

55063

Delaval, Jan

From: Huynh, Phuong N.
Sent: Monday, November 19, 2001 8:12 AM
To: Delaval, Jan
Subject: RE: 09/610,118

Good morning Jan!

Please search polypeptide of SEQ ID NO: 66, 61, 62, 63, 64, 65 against commercial and interference databases.

Thank you ☺

Neon
Art unit 1644
Mail CM1, 9E12
Office CM1, 9D06
Tel 308-4844

264
132
132

Point of Contact:
Jan Delaval
Librarian-Physical Sciences
CM1 1E01 Tel: 308-4498

THIS PAGE BLANK (USPTO)

Access DB#

55063

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Point of Contact:
 Jan Delaval
 Librarian-Physical Sciences
 CM1 1E01 Tel: 308-4498

STAFF USE ONLY

Searcher: Jan
 Searcher Phone #: 2498
 Searcher Location: _____
 Date Searcher Picked Up: 11/19
 Date Completed: 11/19
 Searcher Prep & Review Time: _____
 Clerical Prep Time: 25
 Online Time: 15

Type of Search

NA Sequence (#) _____
 AA Sequence (#) 6
 Structure (#) _____
 Bibliographic 1
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbit _____
 Dr.Link _____
 Lexis/Nexis _____
 Sequence Systems ☒ _____
 WWW/Internet _____
 Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:11:07 ; Search time 48.99 Seconds
(without alignments)
7.775 Million cell updates/sec

Title: US-09-610-118-61

Perfect score: 31

Sequence: 1 SYWMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 segs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	98	2	PL0123
2	31	100.0	101	2	S26460
3	31	100.0	110	2	PH1655
4	31	100.0	114	2	S36280
5	31	100.0	115	2	S38714
6	31	100.0	117	2	S78486
7	31	100.0	117	2	S17079
8	31	100.0	118	2	S00700
9	31	100.0	122	2	A33989
10	31	100.0	122	2	S31675
11	31	100.0	136	2	S31687
12	31	100.0	139	2	I37781
13	31	100.0	147	2	I37780
14	31	100.0	942	2	B84469
15	31	100.0	942	2	T37539
16	31	100.0	1254	1	A32686
17	31	100.0	1844	2	T51890
18	31	100.0	101	2	H37262
19	28	90.3	103	2	PH0986
20	28	90.3	106	2	S26315
21	28	90.3	108	2	S26316
22	28	90.3	110	2	S26317
23	28	90.3	117	2	G45722
24	28	90.3	118	2	C30560
25	28	90.3	120	2	PD0008
26	28	90.3	126	2	E71185
27	28	90.3	161	2	S49488
28	28	90.3	180	2	T39395
29	28	90.3	190	2	A45601

30	28	90.3	190	2	A54518	22.6k tegument ant
31	28	90.3	191	2	S43178	gene s122 protein
32	28	90.3	204	2	B82166	conserved hypotet
33	28	90.3	216	2	A72291	hypothetical prote
34	28	90.3	220	2	B49736	collagen alpha 3(I
35	28	90.3	226	2	T29404	hypothetical prote
36	28	90.3	246	2	I48302	collagen alpha 3(I
37	28	90.3	253	2	I48304	collagen alpha 5(I
38	28	90.3	254	2	B63169	probable transport
39	28	90.3	258	2	B61228	collagen alpha 1(I
40	28	90.3	261	2	A34476	collagen alpha 2(I
41	28	90.3	309	2	S43573	C05B5.4 protein (c
42	28	90.3	312	2	I48303	collagen alpha 4(I
43	28	90.3	342	2	A85635	part of regulation
44	28	90.3	342	2	H64840	part protein precu
45	28	90.3	349	2	C82427	periplasmic protei

ALIGNMENTS

RESULT 1
PL0123
Ig heavy chain V-III region (TD-Vr) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
C:Accession: PL0123; S26897
R:Bird, J.; Gallili, N.; Link, M.; Stiles, D.; Sklar, J.
J. Exp. Med. 168, 229-245, 1988
A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglob
A:Reference number: PL0116; MUID:88286083
A:Accession: PL0123
A:Molecule type: mRNA
A:Residues: 1-98 <BIR>
A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
A:Note: The sequence shows the V region (TD-Vr) from a nonproductive DNA rearrangement
R:Tollinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; MUID:93021117
A:Accession: S26897
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12354; NID:932930; PIDD:CA78224.1; PID:932931
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMH>
F:31-35/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match 100.0%; Score 31; DB 2; Length 98;
Best local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYWMS 5
DB 31 SYWMS 35
RESULT 2
S26460
Ig heavy chain V region - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S26460
R:Kavaler, J.
submitted to the EMBL Data Library, April 1991
A:Reference number: S26459
A:Accession: S26460
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-101 <KAV>
 A:Cross-references: EMBL:X59106; NID:g51707; PIDN:CAA1832.1; PID:g51708
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin

Query Match 100.0%; Score 31; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 13 SYWMS 17

RESULT 3
 PH1655
 Ig heavy chain V region (clone 2E8) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
 C:Accession: PH1655
 R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.
 J. Exp. Med. 178, 331-336, 1993
 A:Title: The structural basis of germ-line encoded VH3 immunoglobulin binding to staphylococcal protein A
 A:Reference number: PH1642; MUID:93301610
 A:Accession: PH1655
 A:Molecule type: mRNA
 A:Residues: 1-110 <HIL>
 A:Experimental source: B cell
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:7-90/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 23 SYWMS 27

RESULT 4
 S36280
 Ig heavy chain V region (clone alpha-FOG1-A3) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
 C:Accession: S36280
 R:Giffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.
 EMBO J. 12, 725-734, 1993
 A:Title: Human anti-self antibodies with high specificity from phage display libraries.
 A:Reference number: S36256; MUID:93178448
 A:Accession: S36280
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-114 <GRI>
 A:Cross-references: EMBL:Z18822
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 114;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 Db 31 SYWMS 35

RESULT 5

S38714
 Ig heavy chain V region - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
 C:Accession: S38714
 R:Cimanis, A.Y.
 Submitted to the EMBL Data Library, November 1993
 A:Reference number: S38713
 A:Accession: S38714
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-115 <CIM>
 A:Cross-references: EMBL:X76014; NID:g416092; PIDN:CAA53601.1; PID:g1334076
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:14-99/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 30 SYWMS 34

RESULT 6
 S78486
 Ig heavy chain V region (clone FL13-28) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 21-Jan-2000
 C:Accession: S78486; S31115
 R:Raaphorst, F.M.
 Submitted to the EMBL Data Library, October 1991
 A:Reference number: S78486
 A:Accession: S78486
 A:Molecule type: mRNA
 A:Residues: 1-117 <RAA>
 A:Cross-references: EMBL:X62965
 A:Experimental source: clone FL13-28
 R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuur
 Eur. J. Immunol. 22, 247-251, 1992
 A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comp
 A:Reference number: S31104; MUID:92111633
 A:Accession: S31115
 A:Molecule type: mRNA
 A:Residues: 1-4, 'L', '6-32', 'G', '34-52', 'E', '54-73', 'K', '75-97 <RAW>
 A:Cross-references: EMBL:X62965
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 117;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 31 SYWMS 35

RESULT 7
 S17079
 Ig heavy chain V-gene (clone HHG19) - human
 C:Species: Homo sapiens (man)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
 C:Accession: S60299; S17079
 R:Kneppers, R.; Fischer, U.; Rajewsky, K.; Gause, A.
 Immunol. Lett. 34, 57-62, 1992
 A:Title: Immunoglobulin heavy and light chain gene sequences of a human CD5 positive
 A:Reference number: S60295; MUID:93122853

A:Accession: S60299
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-117 <KU2>
 A:Cross-references: EMBL:X62128; NID:938340; PIDN:CAA44059.1; PID:938341
 A:Note: the authors did not translate the codons for residues 6, 52, 54, 68, 69, 71, 72,
 C:Genetics:
 A:Introns: 16/1
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 117;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||||
 DB 50 SYMS 54

RESULT 8
 Ig heavy chain V region - African clawed frog
 S00700
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 23-Jul-1999
 C:Accession: S00700
 R:Yamawaki-Kataoka, Y.; Honjo, T.
 Nucleic Acids Res. 15, 5888, 1987
 A:Title: Nucleotide sequences of variable region segments of the immunoglobulin heavy ch
 A:Reference number: S00700; MUID:87289054
 A:Accession: S00700
 A:Molecule type: DNA
 A:Residues: 1-118 <YAM>
 A:Cross-references: EMBL:Y00380; NID:964810; PIDN:CAA68452.1; PID:964811
 A:Note: the sequence was determined from the germline gene
 C:Genetics:
 A:Introns: 15/3
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 118;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||||
 DB 50 SYMS 54

RESULT 9
 A33989
 Ig heavy chain V-1-D-J region - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 16-Aug-1996
 C:Accession: A33989
 R:Hsu, E.; Schwager, J.; Alt, F.W.
 Proc. Natl. Acad. Sci. U.S.A. 86, 8010-8014, 1989
 A:Title: Evolution of immunoglobulin genes: V-H families in the amphibian Xenopus.
 A:Reference number: A33989; MUID:90046727
 A:Accession: A33989
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-122 <HSU>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-100/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 122;

Best Local Similarity 100.0%; Pred. No. 37;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||||
 DB 31 SYMS 35

RESULT 10
 S31675
 Ig heavy chain V region - human
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
 C:Accession: S31675
 R:Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelle, C.
 Submitted to the EMBL Data Library, June 1992
 A:Description: Mechanisms that generate human immunoglobulin diversity operate from t
 A:Reference number: S31585
 A:Accession: S31675
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-122 <CU1>
 A:Cross-references: EMBL:214175; NID:931015; PIDN:CAA78544.1; PID:931016
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:21-104/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||||
 DB 37 SYMS 41

RESULT 11
 S31587
 Ig heavy chain V region - human
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
 C:Accession: S31587
 R:Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelle, C.
 Submitted to the EMBL Data Library, June 1992
 A:Description: Mechanisms that generate human immunoglobulin diversity operate from t
 A:Reference number: S31585
 A:Accession: S31587
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-136 <CU1>
 A:Cross-references: EMBL:214189; NID:931005; PIDN:CAA78558.1; PID:931006
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:31-114/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 136;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||||
 DB 47 SYMS 51

RESULT 12
 I37781
 Ig variable region (VDJ) (clone T21-9) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
 C:Accession: I37781; S25475
 R:Demaison, C.; Chastagner, P.; Theze, J.; Zouali, M.

Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
 A:Title: Somatic diversification in the heavy chain variable region genes expressed by B
 A:Reference number: A36876; MUID:94119917
 A:Accession: 137781
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-139 <RES>
 A:Cross-references: EMBL:X67908; NID:933580; PIDN:CAA48106.1; PID:933581
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 F:35-118/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 139;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 51 SYWMS 55

RESULT 13
 137780
 Ig variable region (VDJ) (clone T20-11) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
 C:Accession: 137780; S23474
 R:Demaison, C.; Chastagner, P.; Theze, J.; Zouali, M.
 Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
 A:Title: Somatic diversification in the heavy chain variable region genes expressed by B
 A:Reference number: A36876; MUID:94119917
 A:Accession: 137780
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-147 <RES>
 A:Cross-references: EMBL:X67943; NID:933578; PIDN:CAA48130.1; PID:933579
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 F:28-111/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 147;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 44 SYWMS 48

RESULT 14
 B84469
 Mutator-like transposase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: B84469
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: B84469
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-942 <STO>
 A:Cross-references: GB:AE002093; NID:94581175; PIDN:AAD24658.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g05490
 A:Map position: 2
 C:Superfamily: Arabidopsis thaliana hypothetical protein F26C24.7

Query Match 100.0%; Score 31; DB 2; Length 942;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 63 SYWMS 67

RESULT 15
 T37539
 Probable guanine nucleotide exchange factor - fission yeast (Schizosaccharomyces pombe
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37539
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21721
 A:Accession: T37539
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-942 <MUR>
 A:Cross-references: EMBL:Z98595; PIDN:CA01190.1; GSPDB:GN00066; SPDB:SPAC11E3.11c
 C:Experimental source: strain 972h-; cosmid c11E3
 C:Genetics:
 A:Gene: SPDB:SPAC11E3.11c
 A:Map position: 1

Query Match 100.0%; Score 31; DB 2; Length 942;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYWMS 5
 Db 539 SYWMS 543

Search completed: November 19, 2001, 08:14:20
 Job time: 193 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:13:32 ; Search time 27.32 Seconds
(without alignments)
6.269 Million cell updates/sec

Title: US-09-610-118-61

Perfect score: 31
Sequence: 1 SYWMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	1254	1 DPOG_YEAST	P15801 saccharomyc
2	31	90.3	180	1 DPOG_SCHPO	O94626 schistosom
3	28	90.3	190	1 TEGU_SCHMA	P14202 schistosoma
4	28	90.3	315	1 YK04 CAEEL	P34292 caenorhabdi
5	28	90.3	342	1 TORT_ECOLI	P38683 escherichia
6	28	90.3	359	1 DNTL_BPT7	P00969 bacterioph
7	28	90.3	385	1 YWBM_BACSU	P39596 bacillus su
8	28	90.3	393	1 Y306_MYCGE	P47548 mycoplasma
9	28	90.3	455	1 YAO6_SCHPO	Q10085 schistosac
10	28	90.3	471	1 CA34_BOVIN	Q28084 bos taurus
11	28	90.3	583	1 PHBC_AZOCA	O66392 a poly-beta
12	28	90.3	745	1 PRL1_RAT	O63517 rattus norv
13	28	90.3	754	1 CA54_CANFA	Q28247 canis fami
14	28	90.3	1669	1 CA14_HUMAN	P02462 homo sapien
15	28	90.3	1669	1 CA14_MOUSE	P02463 mus musculu
16	28	90.3	1670	1 CA34_HUMAN	Q01955 homo sapien
17	28	90.3	1678	1 CA64_HUMAN	Q14031 homo sapien
18	28	90.3	1685	1 CA54_HUMAN	P29400 homo sapien
19	28	90.3	1707	1 CA24_MOUSE	P08122 mus musculu
20	28	90.3	1712	1 CA24_HUMAN	P08572 homo sapien
21	28	90.3	1758	1 CA24_CAEL	P17740 caenorhabdi
22	28	90.3	1763	1 CA24_ASCSU	P27393 ascaris suu
23	28	90.3	2210	1 RRPL_EBOSM	O66802 ebola virus
24	28	90.3	2397	1 MOKB_SCHPO	O09854 schistosac
25	28	90.3	3587	1 SRP2_BACSU	Q00747 bacillus su
26	28	90.3	4540	1 DYRC_PARTE	Q27171 paramecium
27	27	87.1	116	1 HV36_MOUSE	P01806 mus musculu
28	27	87.1	117	1 HV04_MOUSE	P01748 mus musculu
29	27	87.1	117	1 HV05_MOUSE	P01749 mus musculu
30	27	87.1	117	1 HV06_MOUSE	P01750 mus musculu
31	27	87.1	117	1 HV09_MOUSE	P01753 mus musculu
32	27	87.1	117	1 HV10_MOUSE	P01754 mus musculu
33	27	87.1	117	1 HV41_MOUSE	P01811 mus musculu

34	27	87.1	117	1 HV42_MOUSE	P01812 mus musculu
35	27	87.1	117	1 HV49_MOUSE	P06328 mus musculu
36	27	87.1	118	1 HV39_MOUSE	P01809 mus musculu
37	27	87.1	119	1 HV37_MOUSE	P01807 mus musculu
38	27	87.1	119	1 HV38_MOUSE	P01808 mus musculu
39	27	87.1	119	1 HV40_MOUSE	P01810 mus musculu
40	27	87.1	120	1 HV50_MOUSE	P06329 mus musculu
41	27	87.1	139	1 HV07_MOUSE	P01751 mus musculu
42	27	87.1	173	1 NM6M_LATCH	O03175 latimeria c
43	27	87.1	355	1 PYRC_SULAC	O08357 sulfolobus
44	27	87.1	365	1 PD2_DROME	Q02360 drosophila
45	27	87.1	444	1 CIT1_KLEPN	P16482 klebsiella

ALIGNMENTS

```

RESULT 1
DPOG_YEAST STANDARD; PRT; 1254 AA.
ID      DPOG_YEAST
AC      P15801: 008785;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-DEC-1998 (Rel. 37, Last annotation update)
DE      DNA POLYMERASE GAMMA (EC 2.7.7.7) (MITOCHONDRIAL DNA POLYMERASE
DE      CATALYTIC SUBUNIT).
GN      MIP1 OR YOR330C.
OS      Saccharomyces cerevisiae (Baker's Yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX      NCBI_TaxID=4932;
RN      [1]
RX      MEDLINE=90062193; PubMed=2684980;
RA      Foury F.;
RT      "Cloning and sequencing of the nuclear gene MIP1 encoding the
RT      catalytic subunit of the yeast mitochondrial DNA polymerase.";
RL      J. Biol. Chem. 264:20552-20560(1989).
RN      [2]
RX      Foury F.;
RT      REVISIONS.
RL      Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.
RA      [3]
RX      SEQUENCE FROM N.A.
RX      MEDLINE=97051586; PubMed=8896263;
RX      Parle-Mcdermott A.G., Hand N.J., Goulding S.G., Wolfe K.H.;
RT      "Sequence of 29 kb around the PDR10 locus on the right arm of
RT      Saccharomyces cerevisiae chromosome XV: similarity to part of
RT      chromosome I.";
RN      [4]
RX      Yeast 12:999-1004(1996).
RN      [5]
RX      SEQUENCE OF 987-1254 FROM N.A.
RN      RC      STRAIN=GRF88;
RA      Song J.M., Cheung E., Rabinowitz J.C.;
RL      Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: INVOLVED IN THE REPLICATION OF MITOCHONDRIAL DNA.
CC      -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC      N PYROPHOSPHATE + DNA(n).
CC      -1- COFACTOR: MAGNESIUM.
CC      -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC      -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC      ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
CC      DIFFERENT REACTIONS OF DNA SYNTHESIS.
CC      -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----

```

DR EMBL: J05117; AAA17543.1; -
 DR EMBL: 249821; CAA89977.1; -
 DR EMBL: 275238; CAA9652.1; ALT-INT.
 DR EMBL: U42227; AAA85442.1; -
 DR PIR: A32686; A32686.
 DR SGD: S0005857; MIP1.
 DR InterPro: IPR001098; -
 DR InterPro: IPR002297; -
 DR Pfam: PF00476; DNA_POL_A.1.
 DR PRINTS: PRO0667; DNAPOIC.
 DR PROSITE: PS00447; DNA_POLYMERASE_A.1.
 KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Mitochondrion; Magnesium.
 FT DOMAIN 150 189
 FT DOMAIN 198 198
 FT DOMAIN 225 255
 FT DOMAIN 225 255
 FT CONFLICT 8 8
 FT CONFLICT 35 35
 FT CONFLICT 222 222
 FT CONFLICT 357 357
 FT CONFLICT 540 541
 FT CONFLICT 616 616
 FT CONFLICT 661 661
 FT CONFLICT 978 978
 FT CONFLICT 986 986
 SO SEQUENCE 1254 AA; 143501 MW; 80EE686BE9A9F2EC CRC64;

Query Match 100.0%; Score 31; DB 1; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 |||||
 Db 599 SYWMS 603

RESULT 2
 ORN_SCHPO STANDARD; PRT; 180 AA.
 AC 094626;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROBABLE OLIGORIBONUCLEASE (EC 3.1.1.1).
 GN SPBC1347.07.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=972.
 RA Wood V., Rejandream M.A., Barrell B.G., Rieger M.,
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: 3'-TO-5' EXORIBONUCLEASE SPECIFIC FOR SMALL
 CC OLIGORIBONUCLEOTIDES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE OLIGORIBONUCLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL035548; CAB37438.1; -
 KW Hypothetical protein; Hydrolase; Exonuclease; Nuclease.
 FT ACT_SITE 130 130
 FT ACT_SITE 130 130
 SO SEQUENCE 180 AA; 20785 MW; 7AE466905131B603 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 180;
 Best Local Similarity 80.0%; Pred. No. 79;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 |||||
 Db 176 SYWMS 180

RESULT 3
 TEGU_SCHMA STANDARD; PRT; 190 AA.
 AC TEGU_SCHMA
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE TEGUMENT ANTIGEN (1(H)A) (ANTIGEN SMA22.6) (A12).
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigoida; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OX NCBI_TaxID=6183;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=PUERTO RICAN;
 RX MEDLINE=87014570; PubMed=2429181;
 RA Stein L.D., David J.R.;
 RT "Cloning of a developmentally regulated tegument antigen of
 RT Schistosoma mansoni.";
 RL Mol. Biochem. Parasitol. 20:253-264(1986).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91304514; PubMed=1852171;
 RA Jeffs S.A., Hagan P., Allen R., Correa-Oliveira R., Smithers S.R.,
 RA Simpson A.J.G.;
 RT "Molecular cloning and characterisation of the 22-kilodalton adult
 RT protective vaccine with isolated tegumental surface membranes.";
 RL Mol. Biochem. Parasitol. 46:159-168(1991).
 CC -1- TISSUE SPECIFICITY: ADULT TEGUMENT.
 CC -1- DEVELOPMENTAL STAGE: THIS ANTIGEN OCCURS IN ADULTS & SPOROZOYTES
 CC BUT NOT IN CERCARIAE, EGGS OR NEWLY TRANSFORMED SCHISTOSOMULA.
 CC IT IS A DEVELOPMENTALLY REGULATED PROTEIN.
 CC -1- SIMILARITY: TO S.MANSONI ANTIGEN SM21.7.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M29837; AAA29856.1; -
 DR EMBL: M37003; AAA29922.1; -
 DR PIR: A54518; A54518.
 DR HSSP: P07463; 10SA.
 DR InterPro: IPR002048; -
 DR Pfam: PF000036; ehand; 2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 KW Antigen.
 SO SEQUENCE 190 AA; 22578 MW; F265EFA29AA7ACDB CRC64;

Query Match 90.3%; Score 28; DB 1; Length 190;
 Best Local Similarity 80.0%; Pred. No. 83;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 |||||
 Db 160 SYWMS 164

```

RESULT 4
YK04_CAEEL STANDARD; PRT; 315 AA.
AC P34292;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPOTHELTICAL 35.9 KDA PROTEIN C05B5.4 IN CHROMOSOME III.
GN C05B5.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peioderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL NZ;
RA Mortimore B.J.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z32679; CAA83591.1; -
DR WormPep: C05B5.4; CE00880.
KW Hypothetical protein.
SQ SEQUENCE 315 AA; 35918 MW; D665F5FAB040BCF CRC64;

Query Match 90.3%; Score 28; DB 1; Length 315;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
DB 281 TYWMS 285

RESULT 5
TORT_ECOLI STANDARD; PRT; 342 AA.
AC P38683; P75888;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PERIPLASMIC PROTEIN TORT PRECURSOR.
GN TORT.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=94364937; Pubmed=8083154;
RA Simon G., Mejean V., Jourlin C., Chipaux M., Pascal M.-C.;
RA "The tort gene of Escherichia coli encodes a response regulator
RA protein involved in the expression of the trimethylamine N-oxide
RA reductase genes.";
RL J. Bacteriol. 176:5601-5606(1994).
RN [2]
RP REVISIONS TO 29-30, SEQUENCE OF 19-24, AND CHARACTERIZATION.
RC STRAIN-K12;
RX MEDLINE=96165284; Pubmed=8576063;
RA Jourlin C., Simon G., Pommier J., Chipaux M., Mejean V.;
RA "The periplasmic Tort protein is required for trimethylamine N-oxide
RA reductase gene induction in Escherichia coli.";
RL J. Bacteriol. 178:1219-1223(1996).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; Pubmed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:12453-12474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97061202; Pubmed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Horio A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- FUNCTION: UPON BINDING A PUTATIVE INDUCER IT PROBABLY INTERACTS
CC WITH TOES AND ALLOWS IT TO PLAY A ROLE IN THE INDUCTION OF THE
CC TORCAD OPERON FOR TRIMETHYLAMINE N-OXIDE REDUCTASE.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC RECEPTOR FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94231; CAA63921.1; -
DR EMBL: AE000201; AAC74079.1; -
DR EMBL: D90736; BAA36136.1; -
DR EMBL: D90737; BAA35761.1; -
DR Ecogene: EG12616; tort.
DR InterPro: IPR001761; -
DR Pfam: PF00532; Peripla_BP_like. 1.
KW Transport; Periplasmic; Signal.
FT SIGNAL 1 18
FT CHAIN 19 342 PERIPLASMIC PROTEIN TORT.
FT CONFLICT 59 59 W -> S (IN REF. 1).
FT CONFLICT 230 230 A -> R (IN REF. 1).
SQ SEQUENCE 342 AA; 37864 MW; 82FC7AAEDF3B2DB9 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 342;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
DB 57 SYWMS 61

RESULT 6
DNL1_BP7 STANDARD; PRT; 359 AA.
AC P00969;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DNA LIGASE (EC 6.5.1.1) (POLYDEOXYRIBONUCLEOTIDE SYNTHASE [ATP]).
GN 1.3.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae;

```

CC T7-like phages.
 OX NCBI_TaxID=10760;
 RN [1]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=83241725; PubMed=6864790;
 RT Dunn J.J., Studier F.W.;
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
 RT locations of T7 genetic elements.";
 RT J. Mol. Biol. 166:477-535(1983).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=82078034; PubMed=7310871;
 RA Dunn J.J., Studier F.W.;
 RT "Nucleotide sequence from the genetic left end of bacteriophage T7
 RT DNA to the beginning of gene 4.";
 RT J. Mol. Biol. 148:303-330(1981).
 RN [3]
 RN SEQUENCE OF 278-359 FROM N.A.
 RX MEDLINE=81053683; PubMed=7001354;
 RA Dunn J.J., Studier F.W.;
 RT "The transcription termination site at the end of the early region of
 RT bacteriophage T7 DNA.";
 RT Nucleic Acids Res. 8:2119-2132(1980).
 RN [4]
 RN SEQUENCE OF 1-5 FROM N.A.
 RX MEDLINE=81054683; PubMed=6254001;
 RA Saito H., Tabor S., Tamaol F., Richardson C.C.;
 RT "Nucleotide sequence of the primary origin of bacteriophage T7 DNA
 RT replication: relationship to adjacent genes and regulatory
 RT elements.";
 RT Proc. Natl. Acad. Sci. U.S.A. 77:3917-3921(1980).
 RN [5]
 RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=96222302; PubMed=8653795;
 RA Subramanya H.S., Doherty A.J., Ashford S.R., Wigley D.B.;
 RT "Crystal structure of an ATP-dependent DNA ligase from bacteriophage
 RT T7.";
 RT Cell 85:607-615(1996).
 CC -1- FUNCTION: DNA LIGASE, WHICH IS EXPRESSED IN THE EARLY STAGE OF
 CC LYTIC DEVELOPMENT, HAS BEEN IMPLICATED IN T7 DNA SYNTHESIS AND
 CC GENETIC RECOMBINATION. IT MAY ALSO PLAY A ROLE IN T7 DNA REPAIR.
 CC -1- CATALYTIC ACTIVITY: ATP + (DEOXYRIBONUCLEOTIDE)(N) +
 CC (DEOXYRIBONUCLEOTIDE)(N+M).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT DNA LIGASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: V01124; CA24322.1; -;
 DR EMBL: V01126; CA24326.1; -;
 DR EMBL: V01127; CA24336.1; -;
 DR EMBL: V01146; CA24393.1; -;
 DR PIR: A01200; LQBP37;
 DR PIR: S42292; S42292;
 DR PDB: 1A01; 2S-MAR-96;
 DR InterPro: IPR000577; -;
 DR Pfam: PF01068; DNA_LIGASE_1;
 DR PROSITE: PS00697; DNA_LIGASE_A1; 1;
 DR PROSITE: PS00333; DNA_LIGASE_A2; 1;
 DR PROSITE: PS01600; DNA_LIGASE_A3; 1;
 KW DNA repair; DNA replication; DNA recombination; ATP-binding; Ligase;
 KW 3D-structure.
 FT BINDING 34 34 AMP (BY SIMILARITY).
 SO SEQUENCE 359 AA; 4113 MW; 43DA453B71BDF8DC CRC64;

Query Match 90.3% Score 28; DB 1; Length 359;

Best Local Similarity 80.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SYMS 5
 Db 50 SYMLS 54
 RESULT 7
 ID YBWM_BACSU STANDARD; PRT; 385 AA.
 AC P39596;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 42.8 KDA PROTEIN IN EPR-GALK INTERGENIC REGION.
 GN YBWM OR IPA-28D.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95020537; PubMed=7934828;
 RA Glaser P., Kunst F., Arnaut M., Coudart M.P., Gonzales W.,
 RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
 RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
 RA Rapoport G., Danchin A.;
 RT "Bacillus subtilis genome project: cloning and sequencing of the 97
 RT kb region from 325 degrees to 333 degrees.";
 RT Mol. Microbiol. 10:371-384(1993).
 CC -1- SIMILARITY: TO E.COLI YCDO.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X73124; CA51584.1; -;
 DR EMBL: Z99123; CAB15853.1; -;
 DR PIR: S39683; S39683;
 DR Subtilist: BG10574; ywbm.
 KW Hypothetical protein.
 SO SEQUENCE 385 AA; 42796 MW; 1F9DCB6BA385A3C7 CRC64;

Query Match 90.3% Score 28; DB 1; Length 385;
 Best Local Similarity 80.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SYMS 5
 Db 78 SYMLS 82
 RESULT 8
 ID Y306_MYCCE STANDARD; PRT; 393 AA.
 AC P47548;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HYPOTHETICAL PROTEIN MG306.
 GN MG306.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lueder T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RL Science 270:397-403(1995).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39711; AAC71528.1; -.
CC DR TIGR; MG306; -.
CC KW Hypothetical protein; Transmembrane.
CC FT TRANSMEM 15 35 POTENTIAL.
CC FT TRANSMEM 56 76 POTENTIAL.
CC FT TRANSMEM 86 106 POTENTIAL.
CC FT TRANSMEM 131 151 POTENTIAL.
CC FT TRANSMEM 176 196 POTENTIAL.
CC FT TRANSMEM 253 273 POTENTIAL.
CC FT TRANSMEM 289 309 POTENTIAL.
CC FT TRANSMEM 349 369 POTENTIAL.
CC SO SEQUENCE 393 AA; 45750 MW; D740FDA979EC364A CRC64;

Query Match 90.3%; Score 28; DB 1; Length 393;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
DB 173 SYWLS 177

RESULT 9
YA06_SCHPO STANDARD; PRT; 455 AA.
ID YA06_SCHPO
AC 010085;
DC 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHEMETICAL 49.1 KDA PROTEIN C11D9.06 IN CHROMOSOME 1.
GN SPAC11D9.06.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE UPF0013 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL: Z68166; CAA92307.1; -.
CC DR InterPro: IPR002528; -.
CC DR Pfam: PF01554; UPF0013; 1.
CC KW Hypothetical protein; Transmembrane.
CC FT TRANSMEM 4 24 POTENTIAL.
CC FT TRANSMEM 44 64 POTENTIAL.
CC FT TRANSMEM 87 107 POTENTIAL.
CC FT TRANSMEM 159 179 POTENTIAL.
CC FT TRANSMEM 184 204 POTENTIAL.
CC FT TRANSMEM 236 256 POTENTIAL.
CC FT TRANSMEM 280 300 POTENTIAL.
CC FT TRANSMEM 309 329 POTENTIAL.
CC FT TRANSMEM 343 363 POTENTIAL.
CC FT TRANSMEM 381 401 POTENTIAL.
CC FT TRANSMEM 405 425 POTENTIAL.
CC SO SEQUENCE 455 AA; 49086 MW; FF41087BC6405327 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 455;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
DB 196 SYWLS 200

RESULT 10
CA34_BOVIN STANDARD; PRT; 471 AA.
ID CA34_BOVIN
AC Q28084;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE COLLAGEN ALPHA 3(IV) CHAIN (FRAGMENT).
GN COL4A3.
OS Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lens;
RC MEDLINE=91093146; PubMed=1985905;
RA Morrison K.E., Germino G.G., Reeders S.T.;
RA "Use of the polymerase chain reaction to clone and sequence a cDNA
RA encoding the bovine alpha 3 chain of type IV collagen.";
RL J. Biol. Chem. 266:34-39(1991).
RN [2]
RP SEQUENCE OF 227-258.
RC TISSUE=Kidney;
RC MEDLINE=90202779; PubMed=2318822;
RA Gunwar S., Saus J., Noelken M.E., Hudson B.G.;
RA "Glomerular basement membrane. Identification of a fourth chain,
RA alpha 4, of type IV collagen.";
RL J. Biol. Chem. 265:5466-5469(1990).
RN [3]
RP SEQUENCE OF 227-254.
RC MEDLINE=88330844; PubMed=3417661;
RA Saus J., Wieslander J., Langeveld J.P.M., Quinones S., Hudson B.G.;
RA "Identification of the Goodpasture antigen as the alpha 3(IV) chain
RA of collagen IV.";
RL J. Biol. Chem. 263:13374-13380(1988).
RN [4]
RP SEQUENCE OF 227-244.
RC MEDLINE=8722419; PubMed=2438283;
RA Butkowski R.J., Langeveld J.P.M., Wieslander J., Hamilton J.,
RA Hudson B.G.;
RA "Localization of the Goodpasture epitope to a novel chain of basement
RA membrane collagen.";
RL J. Biol. Chem. 262:7874-7877(1987).

```

CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
CC GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
CC MESHWORK TOGETHER WITH LAMININS, PROTEGLYCANS AND ENACTIN/
CC NIDOGEN.
CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC -1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC TRIPLE-HELICAL 7S DOMAIN.
CC -1- PPM: PROLINES AT THE THIRD POSITION OF THE TRIPLEPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- PPM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC ARE INVOLVED IN INTRA- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
CC IV COLLAGENS.
CC -1- SIMILARITY: TO OTHER TYPE IV COLLAGENS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M6139; AAA6708.1; -
DR InterPro: IPR000087; -
DR InterPro: IPR001442; -
DR Pfam: PF01413; C4; 2.
DR Pfam: PF01391; Collagen; 4.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Basement membrane; Collagen; Cell adhesion.
FT NON_TER 1 1
FT DOMAIN <1 238 TRIPLE-HELICAL REGION.
FT SITE 106 108 NONHELICAL REGION (NC1).
FT MOD_RES 232 232 CELL ATTACHMENT SITE (POTENTIAL).
FT MOD_RES 238 238 HYDROXYLATION.
FT DISULFID 261 352 OR 349 (BY SIMILARITY).
FT DISULFID 294 349 OR 352 (BY SIMILARITY).
FT DISULFID 306 312 BY SIMILARITY.
FT DISULFID 371 466 OR 463 (BY SIMILARITY).
FT DISULFID 405 463 OR 466 (BY SIMILARITY).
FT DISULFID 417 423 BY SIMILARITY.
FT CONFLICT 253 253 S -> Y (IN REF. 3).
SQ SEQUENCE 471 AA: 47585 MW: C03B66F1AE7008DE CRC64:

Query Match 90.3%; Score 28; DB 1; Length 471;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYNWS 5
DB 321 SYWLS 325

RESULT 11
PHRC_AZOCA STANDARD; PRT; 583 AA.
AC 066392;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE POLY-BETA-HYDROXYBUTYRATE POLYMERASE (EC 2.3.1.-) (POLY(3-
DE HYDROXYBUTYRATE) POLYMERASE) (PHB POLYMERASE) (POLY(3-
DE HYDROXYALKANATE) POLYMERASE) (PHA SYNTHASE) (POLY(3-
DE (POLYHYDROXYALKANATE) ACID SYNTHASE)).
GN PHRC.

OS Azorhizobium caulinodans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hypnophycozium group; Azorhizobium.
OX NCBI_TaxId-7;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CROSS71;
RA Mandon K., Michel-Reydellet N., Encarnacion S., Kaminski A., Lelja A.,
RA Cavallio M.A., Elmerich C., Mora J.;
RT "Poly-beta-hydroxybutyrate turnover in Azorhizobium caulinodans is
RT required for growth and affects alpha expression."
RT Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
RL -1- FUNCTION: POLYMERIZES D(-)-3-HYDROXYBUTYRYL-COA TO CREATE PHB
CC WHICH CONSISTS OF THOUSANDS OF HYDROXYBUTYRATE MOLECULES LINKED
CC END TO END. PHB SERVES AS AN INTRACELLULAR ENERGY RESERVE MATERIAL
CC WHEN CELLS GROW UNDER CONDITIONS OF NUTRIENT LIMITATION.
CC -1- PATHWAY: THIRD STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PHA/PHB SYNTHASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ006237; CAA06928.1; -
DR InterPro: IPR000073; -
DR Pfam: PF00561; abhydrolase; 1.
KW PHB biosynthesis; transferase; acyltransferase.
FT ACT_SITE 320 320 POTENTIAL.
SQ SEQUENCE 583 AA: 64739 MW: C205763D60DD1A18 CRC64:

Query Match 90.3%; Score 28; DB 1; Length 583;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYNWS 5
DB 50 NYWMS 54

RESULT 12
PRL_RAT STANDARD; PRT; 745 AA.
AC 063517;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE LEUCINE-RICH PRIMARY RESPONSE PROTEIN 1 (FOLLICLE-STIMULATING HORMONE
DE PRIMARY RESPONSE PROTEIN).
GN FSHPR1 OR LRP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId-10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-Testis;
RX MEDLINE-95278605; PubMed-7758824;
RA Slegtenhorst-Begdeman K.E., Post M., Baarends W.M., Thiemmen A.P.M.,
RA Grootegeed J.A.;
RT "Regulation of gene expression in Sertoli cells by
RT follicle-stimulating hormone (FSH): cloning and characterization of
RT LRP1, a primary response gene encoding a leucine-rich protein."
RL Mol. Cell. Endocrinol. 108:115-124(1995).
CC -1- FUNCTION: INVOLVED IN THE RESPONSE OF GONADAL TISSUES TO FOLLICLE-
CC STIMULATING HORMONE.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS, OVARY AND SPLEEN.
CC A MUCH LOWER MRNA LEVEL IS FOUND IN BRAIN AND LUNG, AND NO

```
CC EXPRESSION IS DETECTED IN LIVER, KIDNEY, HEART, MUSCLE, PITUITARY
CC GLAND, PROSTATE, EPIDIDYMIS AND SEMINAL VESICLE.
CC -1- INDUCTION: BY FOLLICLE-STIMULATING HORMONE (FSH).
CC -1- SIMILARITY: SOME, TO S.POMBE MIS6.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X90355; CAA62018.1; -
DR SEQUENCE 745 AA; 85650 MW; 3C9D1AD066D4FDB5 CRC64;
SO
Query Match 90.3%; Score 28; DB 1; Length 745;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYWMS 5
DB 383 SYWMS 387
RESULT 13
CAS4_CANFA STANDARD; PRT: 754 AA.
ID CAS4_CANFA
AC Q28247;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE COLLAGEN ALPHA 5(IV) CHAIN (FRAGMENT).
GN COL4A5.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=SAWOED; TISSUE=Kidney;
RA MEDLINE=94224868; PubMed=81711024;
RT Zheng K., Thorner P.S., Marrano P., Bauman R., McInnes R.R.;
RT "Canine X chromosome-linked hereditary nephritis: a genetic model for
RT human X-linked hereditary nephritis resulting from a single base
RT mutation in the gene encoding the alpha 5 chain of collagen type
RT IV";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3989-3993(1994).
CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
CC GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
CC MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/
CC NIDOGEN.
CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC -1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
CC G-X-Y REPEATS IN THE LONG CENTRAL, TRIPLE-HELICAL DOMAIN (WHICH MAY
CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC TRIPLE-HELICAL 7S DOMAIN.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPLE-HEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC ARE INVOLVED IN INTER- AND INTRA-MOLECULAR DISULFIDE BONDING. 12 OF
CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
CC IV COLLAGENS.
CC -1- DISEASE: A DEFECT IN COL4A5 HAS BEEN FOUND TO BE THE CAUSE OF
CC CANINE X-LINKED HEREDITARY NEPHRITIS (HN), A DISEASE SIMILAR TO
CC THAT IN HUMANS (ALSO REFERRED TO AS ALPORT SYNDROME) CHARACTERIZED
CC BY PROGRESSIVE RENAL FAILURE AND NEUROSENSOR DEAFNESS.
CC -1- SIMILARITY: TO OTHER TYPE IV COLLAGENS.
```

```
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07888; AAB60258.1; -
DR HSSP: Q04656; 2AM0.
DR InterPro: IPR000087; -
DR InterPro: IPR001442; -
DR Pfam: PF01413; C4; 2.
DR Pfam: PF01391; Collagen; 8.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Basement membrane; Collagen; Cell adhesion.
FT NON_TER 1 1
FT DOMAIN <1 530 TRIPLE-HELICAL REGION.
FT DISULFID 531 >754 NONHELICAL REGION (NC1).
FT DISULFID 552 643 OR 640 (BY SIMILARITY).
FT DISULFID 585 640 OR 643 (BY SIMILARITY).
FT DISULFID 597 603 BY SIMILARITY.
FT DISULFID 662 ? OR 754 (BY SIMILARITY).
FT DISULFID 696 754 BY SIMILARITY.
FT DISULFID 708 714 BY SIMILARITY.
FT NON_TER 754 754
FT SEQUENCE 754 AA; 73537 MW; D5E321C287FA925B CRC64;
SO
Query Match 90.3%; Score 28; DB 1; Length 754;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYWMS 5
DB 612 SYWMS 616
RESULT 14
CA14_HUMAN STANDARD; PRT: 1669 AA.
ID CA14_HUMAN
AC P02462;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE COLLAGEN ALPHA 1(IV) CHAIN PRECURSOR.
GN COL4A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=89340433; PubMed=2701944;
RT Sorinen R., Huotari M., Ganguly A., Prockop D.J., Tryggvason K.;
RT "Structural organization of the gene for the alpha 1 chain of human
RT type IV collagen.";
RL J. Biol. Chem. 264:13565-13571(1989).
RN (2)
RP SEQUENCE OF 46-1257 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=88083584; PubMed=3691802;
RX Sorinen R., Haka-Risku T., Prockop D.J., Tryggvason K.;
RT "Complete primary structure of the alpha 1-chain of human basement
RT membrane (type IV) collagen.";
RL FEBS Lett. 225:188-194(1987).
RN (3)
RP SEQUENCE OF 1-943 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=88029471; PubMed=3311751;
RX Brazel D., Oberbauer I., Dieringer H., Babel W., Glanville R.W.,
RX Deutzmann R., Kuehn K.;
```

RT "Completion of the amino acid sequence of the alpha 1 chain of human
RT basement membrane collagen (type IV) reveals 21 non-triplet
RT interruptions located within the collagenous domain."
RL Eur. J. Biochem. 168:529-536(1987).
RN [4]
RP SEQUENCE OF 28-243.
RX MEDLINE=86004708; PubMed=4043082;
RA Glanville R.W., Qian R.O., Siebold B., Ristrell J., Kuehn K.;
RT "Amino acid sequence of the N-terminal aggregation and cross-linking
RT region (75 domain) of the alpha 1 (IV) chain of human basement
RT membrane collagen."
RL Eur. J. Biochem. 152:213-219(1985).
RN [3]
RP SEQUENCE OF 534-1447.
RX MEDLINE=85003629; PubMed=6434307;
RA Babel W., Glanville R.W.;
RT "Structure of human-basement-membrane (type IV) collagen. Complete
RT amino-acid sequence of a 914-residue-long pepsin fragment from the
RT alpha 1(IV) chain."
RL Eur. J. Biochem. 143:545-556(1984).
RN [6]
RP SEQUENCE OF 1256-1669 FROM N.A.
RX MEDLINE=85207819; PubMed=2581963;
RA Phlejanleml T., Trygvasen K., Myers J.C., Kurkinen M., Lebo R.,
RA Cheung M.-C., Prockop D.J., Boyd C.D.;
RT "cDNA clones coding for the pro-alpha1(IV) chain of human type IV
RT procollagen reveal an unusual homology of amino acid sequences in two
RT halves of the carboxyl-terminal domain."
RL J. Biol. Chem. 260:7681-7687(1985).
RN [7]
RP SEQUENCE OF 1259-1669 FROM N.A.
RX MEDLINE=85216555; PubMed=2582422;
RA Binkler J.M., Gudes L.J., Loidl H.R., Wang S.-Y., Rosenbloom J.,
RA Kefalides N.A., Myers J.C.;
RT "Restricted homology between human alpha 1 type IV and other
RT procollagen chains."
RL Proc. Natl. Acad. Sci. U.S.A. 82:3649-3653(1985).
RN [8]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=89034231; PubMed=3182844;
RA Solinen R., Huotari M., Hostilka S.L., Prockop D.J., Trygvasen K.;
RT "The structural genes for alpha 1 and alpha 2 chains of human type IV
RT collagen are divergently encoded on opposite DNA strands and have an
RT overlapping promoter region."
RL J. Biol. Chem. 263:17217-17220(1988).
RN [9]
RP SEQUENCE OF 1441-1669, AND DISULFIDE BONDS.
RX TISSUE-Placenta;
RA MEDLINE=89005112; PubMed=2844531;
RA Siebold B., Deutzmann R., Kuehn K.;
RT "The arrangement of intra- and intermolecular disulfide bonds in the
RT carboxyterminal, non-collagenous aggregation and cross-linking domain
RT of basement-membrane type IV collagen."
RL Eur. J. Biochem. 176:617-624(1988).
RN [1]
RP FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
RP GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
RP MESHWORK TOGETHER WITH LAMININS, PROTEGLYCANS AND ENTACTIN/
RP NIDOGEN.
CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV) -
CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC TRIPLE-HELICAL 7S DOMAIN.
CC -1- PTM: LYSINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT ARE HYDROXYLATED IN ALL CASES AND BIND CARBOHYDRATE.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE

CC IV COLLAGENS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26576; AAA53098.1; -
CC EMBL: J04217; AAA53098.1; JOINED.
CC EMBL: M26550; AAA53098.1; JOINED.
CC EMBL: M26540; AAA53098.1; JOINED.
CC EMBL: M26542; AAA53098.1; JOINED.
CC EMBL: M26543; AAA53098.1; JOINED.
CC EMBL: M26544; AAA53098.1; JOINED.
CC EMBL: M26545; AAA53098.1; JOINED.
CC EMBL: M26546; AAA53098.1; JOINED.
CC EMBL: M26547; AAA53098.1; JOINED.
CC EMBL: M26537; AAA53098.1; JOINED.
CC EMBL: M26538; AAA53098.1; JOINED.
CC EMBL: M26548; AAA53098.1; JOINED.
CC EMBL: M26549; AAA53098.1; JOINED.
CC EMBL: M26551; AAA53098.1; JOINED.
CC EMBL: M26552; AAA53098.1; JOINED.
CC EMBL: M26553; AAA53098.1; JOINED.
CC EMBL: M26554; AAA53098.1; JOINED.
CC EMBL: M26555; AAA53098.1; JOINED.
CC EMBL: M26556; AAA53098.1; JOINED.
CC EMBL: M26557; AAA53098.1; JOINED.
CC EMBL: M26539; AAA53098.1; JOINED.
CC EMBL: M26538; AAA53098.1; JOINED.
CC EMBL: M26559; AAA53098.1; JOINED.
CC EMBL: M26560; AAA53098.1; JOINED.
CC EMBL: M26561; AAA53098.1; JOINED.
CC EMBL: M26562; AAA53098.1; JOINED.
CC EMBL: M26536; AAA53098.1; JOINED.
CC EMBL: M26563; AAA53098.1; JOINED.
CC EMBL: M26541; AAA53098.1; JOINED.
CC EMBL: M26564; AAA53098.1; JOINED.
CC EMBL: M26565; AAA53098.1; JOINED.
CC EMBL: M26566; AAA53098.1; JOINED.
CC EMBL: M26567; AAA53098.1; JOINED.
CC EMBL: M26568; AAA53098.1; JOINED.
CC EMBL: M26569; AAA53098.1; JOINED.
CC EMBL: M26570; AAA53098.1; JOINED.
CC EMBL: M26571; AAA53098.1; JOINED.
CC EMBL: M26572; AAA53098.1; JOINED.
CC EMBL: M26573; AAA53098.1; JOINED.
CC EMBL: M26574; AAA53098.1; JOINED.
CC EMBL: M26575; AAA53098.1; JOINED.
CC EMBL: Y00706; CAA68698.1; -
CC EMBL: X05561; CAA29075.1; -
CC EMBL: M10940; AAA52006.1; -
CC EMBL: M11315; AAA52042.1; -
CC PIR: S16876; CGHU4B.
CC MIM: 120130; -
CC InterPro: IPR000087; -
CC InterPro: IPR001442; -
CC Pfam: PF01413; C4; 2.
CC Pfam: PF01391; Collagen; 21.
CC Extracellular matrix; Connective tissue; Basement membrane;
CC Repeat; Hydroxylation; Glycoprotein; Collagen; Signal.
CC SIGNAL
CC 1 27
CC PROPEP 28 172 AMINO-TERMINAL PROPEPTIDE (7S DOMAIN).
CC CHAIN 173 1669 COLLAGEN ALPHA 1(IV) CHAIN.
CC DOMAIN 173 1440 TRIPLE-HELICAL REGION.
CC FT 1441 1669 NONHELICAL REGION (NC1).
CC CARBOHYD 126 126 N-LINKED (GLCNAC. . .).
CC FT 1460 1551 OR 1548.
CC FT 1493 1548 OR 1551.
CC FT DISULFID 1505 1511

FT DISULFID 1570 1665 OR 1662.
 FT DISULFID 1604 1662 OR 1665.
 FT DISULFID 1616 1622
 FT CONFLICT 237 238 SG -> KE (IN REF. 4).
 FT CONFLICT 241 241 G -> K (IN REF. 4).
 FT CONFLICT 319 319 Q -> A (IN REF. 3).
 FT CONFLICT 719 719 N -> D (IN REF. 5).
 FT CONFLICT 837 837 D -> Y (IN REF. 5).
 FT CONFLICT 842 842 K -> P (IN REF. 5).
 FT CONFLICT 896 896 V -> W (IN REF. 2).
 FT CONFLICT 904 904 E -> Q (IN REF. 5).
 FT CONFLICT 914 914 S -> K (IN REF. 5).
 FT CONFLICT 998 998 S -> K (IN REF. 5).
 FT CONFLICT 1010 1010 K -> P (IN REF. 5).
 FT CONFLICT 1012 1012 S -> K (IN REF. 5).
 FT CONFLICT 1358 1358 E -> Q (IN REF. 5).
 SQ SEQUENCE 1669 AA; 160611 MW; 3BEBADDFEB98BA84 CRC64;

Query Match 90.3% Score 28; DB 1; Length 1669;
 Best Local Similarity 80.0%; Pred. No. 6.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYMS 5
 |||:
 Db 1520 SYMS 1524

RESULT 15
 CA14_MOUSE STANDARD: PRT: 1669 AA.
 ID CA14_MOUSE
 AC P02463:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE COLLAGEN ALPHA 1(IV) CHAIN PRECURSOR.
 GN COL4A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89197932; PubMed-2703490;
 RA Muthukumar G., Blumberg B., Kurkinen M.;
 RT "The complete primary structure for the alpha 1-chain of mouse
 collagen IV. Differential evolution of collagen IV domains.";
 RL J. Biol. Chem. 264:6310-6317(1989).
 RN [2]
 RP SEQUENCE OF 1-1154 FROM N.A.
 RX MEDLINE-88112221; PubMed-3338568;
 RA Wood L., Theriault N., Vogeli G.;
 RT "CDNA clones completing the nucleotide and derived amino acid
 sequence of the alpha 1 chain of basement membrane (type IV) collagen
 from mouse.";
 RL FEBS Lett. 227:5-8(1988).
 RN [3]
 RP SEQUENCE OF 1149-1424 FROM N.A.
 RX MEDLINE-86301886; PubMed-3755692;
 RA Nath P., Laurent M., Horn E., Sobel M.E., Zon G., Vogeli G.;
 RT "Isolation of an alpha 1 type-IV collagen cDNA clone using a
 synthetic oligodeoxynucleotide.";
 RL Gene 43:301-304(1986).
 RN [4]
 RP SEQUENCE OF 1276-1669 FROM N.A.
 RX MEDLINE-85127033; PubMed-2578961;
 RA Oberbauer I., Laurent M., Schwarz U., Sakurai Y., Yamada Y.,
 RA Vogeli G., Voss T., Steboid B., Glanville R.M., Kuhn K.;
 RT "Amino acid sequence of the non-collagenous globular domain (NC1) of
 the alpha 1(IV) chain of basement membrane collagen as derived from
 RT complementary DNA.";
 RL Eur. J. Biochem. 147:217-224(1985).
 RN [5]

RP SEQUENCE OF 1441-1669 FROM N.A.
 RX MEDLINE-87250460; PubMed-3597383;
 RA Kurkinen M., Condon M.R., Blumberg B., Barlow D., Quinones S.,
 RA Saus J., Pihlajaniemi T.;
 RT "Extensive homology between the carboxyl-terminal peptides of mouse
 alpha 1(IV) and alpha 2(IV) collagen.";
 RL J. Biol. Chem. 262:8496-8499(1987).
 RN [6]
 RP PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE-86196099; PubMed-3009468;
 RA Sakurai Y., Sullivan M., Yamada Y.;
 RT "Alpha 1 type IV collagen gene evolved differently from fibrillar
 collagen genes.";
 RL J. Biol. Chem. 261:6654-6657(1986).
 RN [7]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE-89066738; PubMed-3198626;
 RA Kaytes P., Wood L., Theriault N., Kurkinen M., Vogeli G.;
 RT "Head-to-head arrangement of murine type IV collagen genes.";
 RL J. Biol. Chem. 263:19274-19277(1988).
 RN [8]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE-89071759; PubMed-3200851;
 RA Burbelo P.D., Martin G.R., Yamada Y.;
 RT "Alpha 1(IV) and alpha 2(IV) collagen genes are regulated by a
 bidirectional promoter and a shared enhancer.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:9679-9682(1988).
 RN [9]
 RP SEQUENCE OF 1-129 FROM N.A.
 RX MEDLINE-88243724; PubMed-3379041;
 RA Killen P.D., Burbelo P., Sakurai Y., Yamada Y.;
 RT "Structure of the amino-terminal portion of the murine alpha 1(IV)
 collagen chain and the corresponding region of the gene.";
 RL J. Biol. Chem. 263:8706-8709(1988).
 CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
 GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
 MESHWORK TOGETHER WITH LAMININS, PROTEGLYCANS AND ENTACTIN/
 NIDOGEN.
 CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOPFORMS, ALPHA 1(IV) -
 ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
 WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
 CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 TRIPLE-HELICAL 7S DOMAIN.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
 ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 IV COLLAGENS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J03758; AAA37439.1; -;
 DR EMBL: M23333; AAA51625.1; -;
 DR EMBL: J04694; AAA50292.1; -;
 DR EMBL: X06777; CAA29946.1; -;
 DR EMBL: X02201; CAA26132.1; -;
 DR EMBL: M15832; AAA37340.1; -;
 DR EMBL: M14042; AAA37342.1; -;
 DR EMBL: M12879; AAA37343.1; -;
 DR EMBL: M13024; -; NOT_ANNOTATED_CDS.
 DR EMBL: M13025; -; NOT_ANNOTATED_CDS.
 DR EMBL: M13026; AAA37344.1; -;
 DR EMBL: M13027; AAA37345.1; -;

DR EMBL: M13043; AAA37346.1; -.
DR EMBL: J04448; AAA37437.1; -.
DR PIR: A33525; CGMS4B.
DR MGD: MGI:88454; Col4a1.
DR InterPro: IPR000087; -.
DR InterPro: IPR001442; -.
DR Pfam: PF01413; C4; 2.
DR Pfam: PF01391; Collagen; 21.
KW Extracellular matrix; Connective tissue; Basement membrane;
KW Repeat; Hydroxylation; Glycoprotein; Collagen; Signal.
FT SIGNAL 1 27
FT PROPEP 28 172 AMINO-TERMINAL PROPEPTIDE (7S DOMAIN).
FT CHAIN 173 1669 COLLAGEN ALPHA 1(IV) CHAIN.
FT DOMAIN 173 1440 TRIPLE-HELICAL REGION.
FT DOMAIN 1441 1669 NONHELICAL REGION (NC1).
FT DISULFID 1460 1551 OR 1548 (BY SIMILARITY).
FT DISULFID 1493 1548 OR 1551 (BY SIMILARITY).
FT DISULFID 1505 1511 BY SIMILARITY.
FT DISULFID 1570 1665 OR 1662 (BY SIMILARITY).
FT DISULFID 1604 1662 OR 1665 (BY SIMILARITY).
FT DISULFID 1616 1622 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 26 26
FT CONFLICT 186 186 S -> P (IN REF. 2).
FT CONFLICT 319 319 Q -> S (IN REF. 2).
FT CONFLICT 369 369 L -> L (IN REF. 2).
FT CONFLICT 403 403 Q -> F (IN REF. 2).
FT CONFLICT 481 481 P -> L (IN REF. 2).
FT CONFLICT 493 493 Q -> H (IN REF. 2).
FT CONFLICT 712 712 S -> I (IN REF. 2).
FT CONFLICT 813 813 E -> Q (IN REF. 2).
FT CONFLICT 982 982 Q -> H (IN REF. 2).
FT CONFLICT 1397 1397 V -> S (IN REF. 3).
SQ SEQUENCE 1669 AA: 160680 MM: 42916B91E52058E9 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 1669;
Best Local Similarity 80.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYMS 5
Db 1520 SYMLS 1524

Search completed: November 19, 2001, 08:25:33
Job time: 721 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:12:42 ; Search time 81.19 Seconds

(without alignments)
8.148 Million cell updates/sec

Title:
Perfect score:
Sequence:

US-09-610-118-61
31
1 SYWMS 5

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_protent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	31	100.0	95	4	Q9ULB6
2	31	100.0	942	0	Q13690
3	31	100.0	942	10	Q9SL18
4	28	90.3	126	1	059434
5	28	90.3	151	10	Q9PFY3
6	28	90.3	161	11	Q61430
7	28	90.3	179	11	P70165
8	28	90.3	185	5	045206
9	28	90.3	191	5	044372
10	28	90.3	191	5	026511
11	28	90.3	191	5	003528
12	28	90.3	191	5	001372
13	28	90.3	202	6	028772
14	28	90.3	203	6	028682
15	28	90.3	203	6	029032
16	28	90.3	204	2	09KRD0
17	28	90.3	205	6	028274
18	28	90.3	206	8	021462
19	28	90.3	206	8	021463

20	28	90.3	206	8	021464	021464 mytilus tro
21	28	90.3	207	8	021465	021465 mytilus cal
22	28	90.3	210	6	028273	028273 canis fam11
23	28	90.3	212	6	028512	028512 macaca mula
24	28	90.3	212	6	028567	028567 ovls aries
25	28	90.3	216	6	09X0M5	09X0M5 thermotoga
26	28	90.3	225	6	028271	028271 canis fam11
27	28	90.3	226	5	P91057	P91057 caenorhabd1
28	28	90.3	229	4	Q9NYC5	Q9NYC5 homo sapien
29	28	90.3	230	11	063122	063122 rattus norv
30	28	90.3	236	10	09XHV8	09XHV8 oryza sativ
31	28	90.3	245	4	Q9NYC4	Q9NYC4 homo sapien
32	28	90.3	246	11	061435	061435 mus musculu
33	28	90.3	253	11	061436	061436 mus musculu
34	28	90.3	254	2	Q916S2	Q916S2 pseudomonas
35	28	90.3	256	2	Q9KX37	Q9KX37 methylocyst
36	28	90.3	295	5	Q9VIX0	Q9VIX0 drosophila
37	28	90.3	312	11	064457	064457 mus musculu
38	28	90.3	338	4	Q9UJ42	Q9UJ42 homo sapien
39	28	90.3	349	2	Q9KYB1	Q9KYB1 vibr1o chol
40	28	90.3	349	2	Q9KLN3	Q9KLN3 vibr1o chol
41	28	90.3	382	5	Q9N3M5	Q9N3M5 caenorhabd1
42	28	90.3	394	2	P73024	P73024 synechocyst
43	28	90.3	398	10	Q9SDC6	Q9SDC6 oryza sativ
44	28	90.3	414	5	045147	045147 caenorhabd1
45	28	90.3	426	2	Q9PJ00	Q9PJ00 chlamydia m

ALIGNMENTS

RESULT 1
Q9ULB6 PRELIMINARY: PRT; 95 AA.
ID Q9ULB6
AC Q9ULB6:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE IMONOGLOBULIN HEAVY CHAIN (FRAGMENT).
GN VH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tange Y., Kayano H.;
RT "Human VH gene sequence."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -i- SIMILARITY: TO IMONOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.
CC CC
DR EMBL: AB035268; BAA87067.1; -.
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003006; -.
DR InterPro: IPR003596; -.
DR Pfam: PF00047; 19; 1.
DR SMART: SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 95
SQ SEQUENCE 95 AA; 10527 MW; 90A8C6D16D22574A CRC64;

Query Match 100.0%; Score 31; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
Db 30 SYWMS 34

RESULT 2
013690

```

ID 013690 PRELIMINARY; PRT; 942 AA.
AC 013690;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DE 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HYPOTHETICAL 105.9 KDA PROTEIN C1E3.11C IN CHROMOSOME I.
GN SPAC11E3.11C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota: Fungi: Ascomycota: Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: SOME, TO YEAST SEC7.
DR EMBL: Z98595; CAB11190.1; -.
DR InterPro: IPR001849; -.
DR Pfam: PF00169; PH: 1.
DR Prosite: PS00003; PH_DOMAIN: 1.
DR SMART: SM00222; Sec7; 1.
KW Hypothetical protein.
SQ SEQUENCE 942 AA; 105908 MW; A677CE2E619ECDF4 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 3; Length 942;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYMMS 5
DB 539 SYMMS 543

```

```

RESULT 3
Q9SL18 PRELIMINARY; PRT; 942 AA.
AC Q9SL18;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE MURATOR-LIKE TRANSPOSASE.
GN AT2G05490.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta;
OC Magnoliophyta: eudicotyledons: core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-T., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buelin L.A., Ketchum K.A., Lee J.-J., Ruming C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Frazer C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana."
RL Nature 402:761-768(1999).
DR EMBL: AC006220; A024658.1; -.
DR InterPro: IPR000566; -.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN.1.
SQ SEQUENCE 942 AA; 106737 MW; 0DEB3F5F5AC0B71E CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 31; DB 10; Length 942;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYMMS 5
DB 63 SYMMS 67

```

```

RESULT 4
ID 059434 PRELIMINARY; PRT; 126 AA.
AC 059434;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE HYPOTHETICAL 14.0 KDA PROTEIN PH1762.
GN PH1762.
OS Pyrococcus horikoshii.
OC Archaea: Euryarchaeota: Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=33953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
DR EMBL: AP000007; BAA30876.1; -.
KW Hypothetical protein.
SQ SEQUENCE 126 AA; 14033 MW; 5FCE589EBA82F83 CRC64;

```

```

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 126;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYMMS 5
DB 69 SYMMS 73

```

```

RESULT 5
ID 09FFY3 PRELIMINARY; PRT; 151 AA.
AC 09FFY3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE GENOMIC DNA, CHROMOSOME 5, P1 CLONE:MNC12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta;
OC Magnoliophyta: eudicotyledons: core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT P1 clones."
RL DNA Res. 4:215-230(1997).
DR EMBL: AB005230; BAB11118.1; -.
SQ SEQUENCE 151 AA; 16532 MW; 52A590132FDB1BC CRC64;

```

Query Match 90.3%; Score 28; DB 10; Length 151;
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 DB 130 SYWMS 134

RESULT 6
 ID 061430 PRELIMINARY; PRT: 161 AA.
 AC 061430;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 DE COLLAGEN TYPE IV ALPHA3 CHAIN (FRAGMENT).
 GN COL4A3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Oberhaeumer I.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X82205; CA57689.1; -.
 DR InterPro: IPR000504; -.
 DR InterPro: IPR001442; -.
 DR Pfam: PF01413; C4; 2.
 DR PROSITE: PS00030; RNP_1; UNKNOWN_1.
 DR SMART: SM00111; C4; 1.
 FT NON_TER 1 1
 FT NON_TER 161 161
 SQ SEQUENCE 161 AA; 17925 MW; 1f59DF6CFE8236C5 CRC64;

Query Match 90.3%; Score 28; DB 11; Length 161;
 Best Local Similarity 80.0%; Pred. No. 3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 DB 29 SYWMS 33

RESULT 7
 ID P70165 PRELIMINARY; PRT: 179 AA.
 AC P70165;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 DE COLLAGEN TYPE IV ALPHA3 CHAIN (FRAGMENT).
 GN COL4A5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Oberhaeumer I.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X82218; CA57698.1; -.
 DR InterPro: IPR001442; -.
 DR Pfam: PF01413; C4; 2.
 DR SMART: SM00111; C4; 1.
 FT NON_TER 1 1
 FT NON_TER 179 179
 SQ SEQUENCE 179 AA; 19859 MW; 20A18BF3687F582F CRC64;

Query Match 90.3%; Score 28; DB 11; Length 179;
 Best Local Similarity 80.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 DB 57 SYWMS 61

RESULT 8
 ID 045206 PRELIMINARY; PRT: 185 AA.
 AC 045206;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 DE 21.7 KDA ANTIGEN.
 GN SJ21.7.
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OC NCBI_TaxID=6182;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PHILIPPINE (MINDORO);
 RA Hafalla J.C., Alamares J.A., Acosta L.P., Ramirez B.L., Santlago M.L.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO EF-HAND FAMILY.
 DR EMBL: AF048759; AAD13338.1; -.
 DR InterPro: IPR002048; -.
 DR Pfam: PF00036; ehand; 2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 DR KW Calcium-binding.
 SQ SEQUENCE 185 AA; 21693 MW; 67D8CCDD2EB84B5 CRC64;

Query Match 90.3%; Score 28; DB 5; Length 185;
 Best Local Similarity 80.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
 DB 155 SYWMS 159

RESULT 9
 ID 044372 PRELIMINARY; PRT: 191 AA.
 AC 044372;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 DE 22.6 KDA TEGUMENTAL ANTIGEN.
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OC NCBI_TaxID=6182;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MINDORO; Hafalla J.C.R., Kurtis J.D., Wiest P.M., Olds G.R.,
 RA Santlago M.L., Ramirez B.L.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO EF-HAND FAMILY.
 DR EMBL: AF030404; AAC67308.1; -.
 DR InterPro: IPR002048; -.
 DR Pfam: PF00036; ehand; 2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 DR SMART: SM00054; EFh; 1.
 KW Calcium-binding.

SO SEQUENCE 191 AA; 22546 MW; 4B163F2D56F97E42 CRC64;

Query Match

Best Local Similarity 90.3%; Score 28; DB 5; Length 191;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYNMS 5

DB 161 SYMN 165

RESULT 10

ID 026511

AC 026511 PRELIMINARY; PRT; 191 AA.

DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)

GN S322.6

OS Schistosoma japonicum (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;

OC Trematoda; Digenea; Strigoidae; Schistosomatidae; Schistosomatidae;

OC Schistosoma.

OX NCBI_TaxID=6182;

RP SEQUENCE FROM N.A.

RC STRAIN-CHINESE MAINLAND ISOLATE;

RA Merckelbach A., Eschbeck A., Ruppel A.;

RL Submitted (Apr-1994) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: TO EF-HAND FAMILY.

DR EMBL: X78574; CAA55312.1; -

DR InterPro: IPR002048; -

DR Pfam: PF00036; efhand; 2.

DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.

DR SMART: SM00054; EFh; 1.

KW Calcium-binding; Signal.

FT SIGNAL 1

FT SIGNAL 17

SO SEQUENCE 191 AA; 22544 MW; 8EP83011A535535D CRC64;

Query Match

Best Local Similarity 90.3%; Score 28; DB 5; Length 191;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYNMS 5

DB 161 SYMN 165

RESULT 11

ID 003528

AC 003528 PRELIMINARY; PRT; 191 AA.

DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)

GN 22.6 KDA TEGUMENT ANTIGEN.

OS Schistosoma japonicum (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;

OC Trematoda; Digenea; Strigoidae; Schistosomatidae; Schistosomatidae;

OC Schistosoma.

OX NCBI_TaxID=6182;

RP SEQUENCE FROM N.A.

RC MEDLINE-94252577; PubMed-8194761;

RA Maine G.J., Becker M., Scott J.C., Kallina B., Yang W., Mcmanus D.P.;

RL "Purification of a recombinant Schistosoma japonicum antigen homologous to the 22-kDa membrane-associated antigen of S. mansoni, a putative vaccine candidate against schistosomiasis.";

CC -1- TISSUE SPECIFICITY: ADULT TEGUMENT.

-1- DEVELOPMENTAL STAGE: THIS ANTIGEN OCCURS IN ADULTS & SPOROCYSTS

BUT NOT IN CERCARIA, EGGS OR NEWLY TRANSFORMED SCHISTOSOMULA.

DR EMBL: I08198; AAA29940.1; -

DR InterPro: IPR002048; -

DR Pfam: PF00036; efhand; 2.

DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.

DR SMART: SM00054; EFh; 1.

KW Antigen.

SO SEQUENCE 191 AA; 22588 MW; 9AFEEA11A5288282 CRC64;

Query Match

Best Local Similarity 90.3%; Score 28; DB 5; Length 191;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYNMS 5

DB 161 SYMN 165

RESULT 12

ID 001372

AC 001372 PRELIMINARY; PRT; 191 AA.

DT 01-JUL-1997 (TREMblrel. 04, Created)

DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)

DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)

GN 22.6 KDA MEMBRANE-ASSOCIATED ANTIGEN.

OS Schistosoma japonicum (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;

OC Trematoda; Digenea; Strigoidae; Schistosomatidae; Schistosomatidae;

OC Schistosoma.

OX NCBI_TaxID=6182;

RP SEQUENCE FROM N.A.

RC STRAIN-JAPANESE;

RA Chen L., Gao P., Yamashita T., Nara T., Kojima S., Senda F., Araki Y.;

RL "Molecular cloning and expression of a tegumental membrane-associated antigen from Japanese strain Schistosoma japonicum.";

CC -1- SIMILARITY: TO EF-HAND FAMILY.

DR EMBL: U75510; AAB52407.1; -

DR EMBL: AB006459; BAA31753.1; -

DR InterPro: IPR002048; -

DR Pfam: PF00036; efhand; 2.

DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.

DR SMART: SM00054; EFh; 1.

KW Calcium-binding.

SO SEQUENCE 191 AA; 22574 MW; 3AE5F011A5289354 CRC64;

Query Match

Best Local Similarity 90.3%; Score 28; DB 5; Length 191;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYNMS 5

DB 161 SYMN 165

Db 161 SYMM 165

RESULT 13

028272 ID Q28272 PRELIMINARY; PRT: 202 AA.
 AC Q28272;
 DT 01-NOV-1996 (TREMBLREL. 01, Created)
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
 DE COLLAGEN TYPE IV ALPHA 2 CHAIN (FRAGMENT).
 GN COL4A2.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SAMVED;
 RX MEDLINE=96278820; PubMed=8662866;
 RA Thorner P.S., Zheng K., Kalluri R., Jacobs R., Hudson B.G.;
 RT "Coordinate gene expression of the alpha3, alpha4, and alpha5 chains
 of collagen type IV. Evidence from a canine model of X-linked
 nephritis with a COL4A3 gene mutation."
 RL J. Biol. Chem. 271:13821-13828(1996).
 DR EMBL: U50934; AAC48584.1; -
 DR InterPro: IPR001442; -
 DR Pfam: PF01413; C4; 2.
 DR SMART: SM00111; C4; 1.
 FT NON_TER 1
 FT NON_TER 202
 SQ SEQUENCE 202 AA; 22079 MW; 25A56E7642A329FC CRC64;

Query Match 90.3%; Score 28; DB 6; Length 202;
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||:1
 Db 70 SYMS 74

RESULT 14

028682 ID Q28682 PRELIMINARY; PRT: 203 AA.
 AC Q28682;
 DT 01-NOV-1996 (TREMBLREL. 01, Created)
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
 DE ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).
 GN COL4A3
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-RENAL CORTEX;
 RA Turner A.N., Ryan J.J., Derry C.J., Cashman S.J., Katbama I.,
 RA Mason P.J., Pusey C.D.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L47283; AAA91893.1; -
 DR InterPro: IPR000504; -
 DR InterPro: IPR001442; -
 DR Pfam: PF01413; C4; 2.
 DR PROSITE: PS00030; RNP_1; UNKNOWN_1.
 DR SMART: SM00111; C4; 1.
 FT NON_TER 1
 FT NON_TER 203
 SQ SEQUENCE 203 AA; 22213 MW; E14173816E4D9E30 CRC64;

Query Match 90.3%; Score 28; DB 6; Length 203;
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||:1
 Db 62 SYMS 66

RESULT 15

029032 ID Q29032 PRELIMINARY; PRT: 203 AA.
 AC Q29032;
 DT 01-NOV-1996 (TREMBLREL. 01, Created)
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
 DE ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).
 GN COL4A3.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-RENAL CORTEX;
 RA Turner A.N., Ryan J.J., Derry C.J., Cashman S.J., Katbama I.,
 RA Mason P.J., Pusey C.D.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L47284; AAA91882.1; -
 DR InterPro: IPR000504; -
 DR InterPro: IPR001442; -
 DR Pfam: PF01413; C4; 2.
 DR PROSITE: PS00030; RNP_1; UNKNOWN_1.
 DR SMART: SM00111; C4; 1.
 FT NON_TER 1
 FT NON_TER 203
 SQ SEQUENCE 203 AA; 22326 MW; E3B20E32D9A245AA CRC64;

Query Match 90.3%; Score 28; DB 6; Length 203;
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
 |||:1
 Db 62 SYMS 66

Search completed: November 19, 2001, 08:25:00
 Job time: 738 sec.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:11:07 ; Search time 77.92 Seconds
(without alignments)
3.890 Million cell updates/sec

Title: US-09-610-118-61
Perfect score: 31
Sequence: 1 SYMMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq-0601.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	5	22	AA61293 Anti-TANCO 268 scf
2	31	100.0	98	15	AA62066 Heavy chain variab
3	31	100.0	98	16	AA62074 DP54 VH region. H
4	31	100.0	98	21	AA640138 Anti-hlil2 antibod
5	31	100.0	98	21	AA640140 Anti-hlil2 antibod
6	31	100.0	117	16	AA66301 Human immunoglobul
7	31	100.0	123	22	AA662777 Human HIV-1 monocl
8	31	100.0	123	22	AA662778 Human HIV-1 monocl
9	31	100.0	125	22	AA662776 Human HIV-1 monocl
10	31	100.0	271	20	AA690180 Human clone A6 fus
11	31	100.0	312	20	AA693323 Single chain Apo-2

12	28	90.3	5	16	AA65499
13	28	90.3	5	17	AA67316
14	28	90.3	5	20	AA69038
15	28	90.3	59	21	AA64781
16	28	90.3	115	21	AA612172
17	28	90.3	118	18	AA601579
18	28	90.3	118	18	AA601576
19	28	90.3	119	17	AA697334
20	28	90.3	119	18	AA601578
21	28	90.3	121	14	AA641655
22	28	90.3	131	15	AA63128
23	28	90.3	140	14	AA64169
24	28	90.3	140	16	AA69480
25	28	90.3	159	21	AA641853
26	28	90.3	190	11	AA603988
27	28	90.3	211	21	AA695918
28	28	90.3	218	16	AA679164
29	28	90.3	218	20	AA644172
30	28	90.3	218	21	AA656784
31	28	90.3	229	10	AA693524
32	28	90.3	229	21	AA67943
33	28	90.3	240	16	AA65495
34	28	90.3	241	13	AA626981
35	28	90.3	241	18	AA609435
36	28	90.3	242	18	AA609444
37	28	90.3	242	19	AA63881
38	28	90.3	242	21	AA67946
39	28	90.3	245	21	AA67942
40	28	90.3	246	18	AA609434
41	28	90.3	247	18	AA609443
42	28	90.3	253	22	AA676513
43	28	90.3	258	20	AA612992
44	28	90.3	258	21	AA67954
45	28	90.3	260	20	AA631991

ALIGNMENTS

RESULT 1	AA61293	standard; Peptide: 5 AA.
ID	AA61293	
XX	AA61293:	
XX	04-APR-2001 (first entry)	
XX	Anti-TANCO 268 scfV CDR, SEQ ID NO: 61.	
DE	Human; antibody; scfV; CDR; complementarity determining region;	SCfV(FWP51) CDRH.
KW	TANCO 268; caridant; cerebroprotective; cyostatic; anticoagulant;	Humanised monoclon
KW	thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;	Tumour antigen ant
KW	platelet membrane glycoprotein receptor; bleeding disorder;	Human 5' EST relat
KW	blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;	Humanised HBV pre-
KW	leukemia; cardiovascular disease; immunological disease; liver disorder;	Lead binding Mab 1
KW	cancer.	Lead binding Mab 6
XX		Murine NEMW MN-14
OS	Homo sapiens.	Lead binding Mab 8
XX		Mab 25D2 VH. Rat
PN	WO200100810-A1.	Anti-interleukin-4
XX		Mab 25D2 humanised
PD	04-JAN-2001.	Human IL-4 hu
XX		Human OREX ORF1617
PF	30-JUN-2000; 2000MO-US18152.	Mr 22k peripheral
XX		Human Goodpasture
XX	30-JUN-1999; 99US-0345468.	Partial sequence o
PR	06-DEC-1999; 99US-0454824.	Human type IV coll
XX	14-FEB-2000; 2000US-0503387.	Human alpha3 type
XX		Complete sequence
PA	(MILL-) MILLENNIUM PHARM INC.	Human type IV coll
XX		ScfV(FWP51). Synt
PI	Busfield ST. Villelaj J. Jandroct-Perrus M. Vainchencker W. Gill DS.	FWP51 fusion prote
PI	Qian MD, Kingsbury G.	Anti-CD19 antibody
		Modified single ch
		Cornebacterium q1
		Type IV collagen N
		Type IV collagen N

XX WPI: 2001-080877/09.
 XX
 PT New genes encoding human platelet-expressed collagen receptor.
 PT glycoprotein VI, and its modulators, useful for preventing, treating
 PT and diagnosing hemorrhagic disorders, thrombotic diseases and
 PT immunological disorders.
 XX
 PS Claim 31: Page 102; 227pp; English.

CC The present sequence is given in a specification relating to an isolated
 CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
 CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
 CC and polypeptides and their modulators, e.g. antisense nucleic acids,
 CC ribozymes and antibodies, are useful for preventing, treating and
 CC diagnosing disorders associated with aberrant expression or activity of
 CC GPVI. These disorders include bleeding disorders
 CC (e.g. thrombocytopaenia), blood vessel injury, thrombotic disorders
 CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
 CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
 CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
 CC infarction), immunological diseases (e.g. platelet disorder) and
 CC embryonic liver disorders. Preferably they are used to prevent acute
 CC cardiac ischaemia following angioplasty and metastatic cancers,
 CC especially of the colon and liver.

CC Sequence 5 AA;

Query Match 100.0%; Score 31; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMMS 5
 |||||
 Db 1 symms 5

RESULT 2

AAR52066
 ID AAR52066 standard; Protein; 98 AA.

AC AAR52066;

DT 11-OCT-1996 (first entry)

DE Heavy chain variable region of human PL0123 antibody.

KW antibody; humanised; murine; human; heavy chain; light; variable;

KW framework region; complementarity determining region; reshaping;

OS Homo sapiens.

Key Location/Qualifiers

FT Region 1..30

FT /label= framework_region_1

FT /note= "FR 1"

FT Region 31..35

FT /label= complementarity_determining_region_1

FT /note= "CDR 1"

FT Region 36..49

FT /note= "FR 2"

FT Region 50..59

FT /note= "CDR 2"

FT Region 60..98

FT /note= "FR 3"

EP592106-A1.

13-APR-1994.

07-SEP-1993; 93EP-0307051.

XX 09-SEP-1992; 9205-0942245.

PA (PEDE/) PEDERSEN J T.
 PA (IMMU-) IMMUNOGEN INC.

PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
 WPI: 1994-120230/15.

PT Method of resurfacing of rodent antibodies to produce humanised
 PT antibody forms - for producing non-human antibodies with improved
 PT therapeutic efficiency by presenting human surface on V-region

PS Example 1: Fig 4B; 230pp; English.

CC Modification of a rodent antibody (Ab) or fragment by resurfacing in
 CC order to produce a humanised rodent Ab can be determined by calculating
 CC homology between murine and human Ab antibody surfaces. In order to test
 CC the resurfacing approach of the invention, three humanisation
 CC experiments were set up: (1) traditional loop grafting; (2) resurfacing
 CC approach using most similar chain; and (3) resurfacing approach using
 CC human sequences with most similar surface residues. The Ab used was the
 CC murine anti-N901 Ab (see AAR52061). Experiment 3 was carried out using
 CC the present sequence which represents the human PL0123 Ab light chain
 CC (AAR52067) was prep. by resurfacing. Sequence numbering starts at
 CC position 118 in the specification.

CC Sequence 98 AA;

Query Match 100.0%; Score 31; DB 15; Length 98;
 Best Local Similarity 100.0%; Pred. No. 51;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMMS 5
 |||||
 Db 31 symms 35

RESULT 3

AAR72074
 ID AAR72074 standard; Protein; 98 AA.

AC AAR72074;

DT 26-SEP-1995 (first entry)

DE DP54 VH region.

KW Graves ophthalmopathy associated immunoglobulin protein;

KW orbital antigen; monoclonal antibody; heavy chain; H chain;

KW variable region; autoimmunity.

Key Location/Qualifiers

FT Region 41..44

FT /label= CDR1

FT Region 49..67

FT /label= CDR2

PN WO9508336-A.

PD 30-MAR-1995.

PF 22-SEP-1994; 94WO-US10756.

PR 22-SEP-1993; 9305-0124469.

PA (NICH-) NICHOLS INST DIAGNOSTICS.

PI McLachlan SM, Rapoport B:
XX WPI: 1995-139383/18.
DR N-PSDB: AA089332.
XX
PT Graves' ophthalmopathy-associated monoclonal antibody - produced
PR by molecular cloning of immunoglobulin genes by PCR
XX
PS Disclosure: Page 72: 94pp; English.
XX
CC L- and H-chain DNA was amplified by PCR from Graves' orbital
CC tissue and clones encoding autoimmune-associated immunoglobulin
CC fragments were obtained. 2/15 clones of H chain (1951) genes
CC showed homology to the germ-line gene DP54 (AA089332), which encodes
CC the protein given in AAR72074. The DNA (AA089333) and corresp. amino
CC acid (AAR2075) sequences of the VH region of a representative clone,
CC OFH1.19, are provided.
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 31; DB 16; Length 98;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
11111
DB 31 sywms 35

RESULT 4
AAB40138
ID AAB40138 standard; Protein; 98 AA.
XX
AC AAB40138;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 664.
XX
KW Human; neutralising antibody; Interleukin-12; IL-12; antiinflammatory;
KW complementarily determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PF 24-MAR-2000: 2000MO-US07946.
XX
PR 25-MAR-1999: 99US-0126603.
XX
PA (BADI) BASF AG.
XX (GEMY) GENETICS INST INC.
XX
PI Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
XX
DR WPI: 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
XX
PS Claim 75: Page 122; 377pp; English.
XX
CC This invention relates to a new human antibody specific for human

CC Interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarily determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
CC represent anti-IL-12 CDR3 related amino acid sequences. AAB39772-B40063
CC given in AAB40064-B40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in
CC AAC61062-C61071. The antibody of the invention is a neutralising
CC antibody and has antirheumatic; antiarthritic; antisclerotic;
CC antiinflammatory; neuroprotective; antipsoriatic; antiasthmatic;
CC cardiant; antiparasitic; antibacterial and immunosuppressive activity.
CC The antibodies or antigen-binding fragments are useful in the treatment
CC of disorders associated with detrimental release of human IL-12,
CC especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
CC They can also be used in the manufacture of a pharmaceutical composition
CC to treat human IL-12 disorders.
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 31; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
11111
DB 31 sywms 35

RESULT 5
AAB40140
ID AAB40140 standard; Protein; 98 AA.
XX
AC AAB40140;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 666.
XX
KW Human; neutralising antibody; Interleukin-12; IL-12; antiinflammatory;
KW complementarily determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PF 24-MAR-2000: 2000MO-US07946.
XX
PR 25-MAR-1999: 99US-0126603.
XX
PA (BADI) BASF AG.
XX (GEMY) GENETICS INST INC.
XX
PI Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
XX
DR WPI: 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
XX
PS Claim 75; Page 122; 377pp; English.

```

Query Match      100.0%; Score 31; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
Qy      1 SYWMS 5
        |||||
Db      31 sywms 35

```

CC	amplification using primers AAQ78917-38. The genes are subdivided into 5
CC	families of Vn genes. The fragments cover a region of 800 kb. The DNA
CC	fragments were isolated from high molecular weight DNA from human
CC	placenta. The DNA was partially digested with XbaI restriction enzyme.
CC	The fragments were separated by gel electrophoresis and 35-45 kb fractions
CC	were collected. The fragments were ligated with ClaI-digested cosmid
CC	vector pJ8B1. The ligation products were in vitro packed and infected
CC	into E.coli 450A. The fragments were then subcloned by colony
CC	hybridisation. The Vn genes and the DNA fragments encoding them are
CC	used in producing human immunoglobulin in mammalian hosts.
XX	
SQ	Sequence 117 AA;
Query Match	100.0%; Score 31; DB 16; Length 117;
Best Local Similarity	100.0%; Pred. No. 61;
Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 SYWMS 5
Db	50 sywms 54
RESULT 7	
AAB62777	'
ID AAB62777	standard; Protein: 123 AA.
XX AC	
XX AAB62777:	
DT 03-APR-2001	(first entry)
DE Human HIV-1 monoclonal antibody SEQ ID NO: 76.	
XX XX	
KM Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;	
KV envelope glycoprotein; gp120; diagnosis.	
XX OS	
OS Homo sapiens.	
XX PN	
PN WO200100678-A1.	
PD 04-JAN-2001.	
XX PF	
PF 23-JUN-2000; 2000WO-US17327.	
XX PR	
PR 30-JUN-1999; 99US-0141701.	
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX PI	
PI Watkins BA, Reitz MS;	
XX DR	
DR WPI: 2001-112438/12.	
XX N-PSDB; AAF29078.	
XX PT	
PT Novel human monoclonal antibody immunoreactive with human	
PT immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1	
PT in biological sample and providing passive immunotherapy to HIV-1	
PT infected mammal	
XX PS	
PS Claim 1; Page 70; 81pp; English.	
XX CC	
CC The present invention provides the protein and coding sequences for the	
CC variable regions of human monoclonal antibodies which are immunoreactive	
CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.	
CC These can be used in diagnosis and therapy of HIV-1 infection.	
XX SQ	
SQ Sequence 123 AA;	
Query Match	100.0%; Score 31; DB 22; Length 123;
Best Local Similarity	100.0%; Pred. No. 64;
Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 SYWMS 5

Db 32 syms 36

RESULT 8

ID AAB62778 standard; Protein: 123 AA.

AC AAB62778;

DT 03-APR-2001 (first entry)

DE Human HIV-1 monoclonal antibody SEQ ID NO: 77.

DE Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;

KW envelope glycoprotein; gp120; diagnosis.

OS Homo sapiens.

PN WO200100678-A1.

PD 04-JAN-2001.

PF 23-JUN-2000; 2000WO-US17327.

PR 30-JUN-1999; 99US-0141701.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Watkins BA, Reltz MS;

DR WPI: 2001-112438/12.

DR N-PSDB: AAF29079.

XX Novel human monoclonal antibody immunoreactive with human

PT immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1

PT in biological sample and providing passive immunotherapy to HIV-1

PT infected mammal

PS Claim 1; Page 71; 81pp; English.

CC The present invention provides the protein and coding sequences for the

CC variable regions of human monoclonal antibodies which are immunoreactive

CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.

CC These can be used in diagnosis and therapy of HIV-1 infection.

CC Sequence 123 AA;

QY 1 SYMS 5

DB 32 syms 36

RESULT 9

ID AAB62776 standard; Protein: 125 AA.

AC AAB62776;

DT 03-APR-2001 (first entry)

DE Human HIV-1 monoclonal antibody SEQ ID NO: 75.

DE Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;

KW envelope glycoprotein; gp120; diagnosis.

OS Homo sapiens.

PN WO200100678-A1.

PD 04-JAN-2001.

PF 23-JUN-2000; 2000WO-US17327.

PR 30-JUN-1999; 99US-0141701.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Watkins BA, Reltz MS;

DR WPI: 2001-112438/12.

DR N-PSDB: AAF29077.

XX Novel human monoclonal antibody immunoreactive with human

PT immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1

PT in biological sample and providing passive immunotherapy to HIV-1

PT infected mammal

PS Claim 1; Page 69-70; 81pp; English.

CC The present invention provides the protein and coding sequences for the

CC variable regions of human monoclonal antibodies which are immunoreactive

CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.

CC These can be used in diagnosis and therapy of HIV-1 infection.

QY 1 SYMS 5

DB 32 syms 36

RESULT 10

ID AAM90180 standard; Protein: 271 AA.

AC AAM90180;

DT 10-MAY-1999 (first entry)

DE Human clone A6 fusion protein antibody.

DE Antibody; human; clone A6; hab; fusion protein; polyclonal tag;

KW Western blot; enzyme-linked immunosorbent assay; ELISA; therapy;

KW immunofluorescence; immunoprecipitation assay; affinity purification;

KW diagnosis; vaccine; serum; immune response.

OS Homo sapiens.

PN WO9901475-A2.

PD 14-JAN-1999.

PF 03-JUL-1998; 98WO-DE01882.

PR 04-JUL-1997; 97DE-1028697.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

PI Braunagel M, Doersam H, Kipriyanov S, Kuerschner T;

PI Little M, Weischof M;

XX

DR WPI: 1999-106000/09.
DR N-PSDB: AAV74274.
XX Human antibody against fusion protein with polyhistidine tag -
PT useful as standard in immunassays, for affinity purification,
PT diagnosis and therapy and for preparing vaccines
XX
PS Claim 2; Fig 4; 20pp; German.
XX
XX This invention describes a human antibody (hAb) against a fusion
CC (poly)peptide or protein that includes a segment of at least 6
CC consecutive His residues. This antibody is useful in Western blots,
CC enzyme-linked immunosorbent assay (ELISA), immunofluorescence or
CC immunoprecipitation assays. Also hAb can be used for affinity
CC purification of the protein, for in vivo diagnosis or therapy, and
CC in production of vaccines. hAb are universally applicable
CC alternatives to human serum. They are specific for polyhistidine
CC tag, regardless of the nature of the rest of the protein. Since hAb
CC are not produced in an animal, they contain no components that can
CC induce an immune response in humans.
SO Sequence 271 AA;

Query Match 100.0%; Score 31; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWMS 5
DB 33 sywms 37

RESULT 11
AAW83323
ID AAW83323 standard; Protein: 312 AA.
XX
AC AAW83323;
XX
DT 16-MAR-1999 (first entry)
XX
DE Single chain Apo-2 antibody 20E6.
XX
XX Human; Apo-2; receptor; apoptosis; neurodegenerative disease; cancer;
KW tumour necrosis factor; TNF; tumour necrosis factor receptor; TNFR;
KW TNF cytokine.
XX
OS Homo sapiens.
XX
PN WO951793-A1.
XX
PD 19-NOV-1998.
XX
PF 14-MAY-1998; 98WO-US09704.
XX
PR 09-FEB-1998; 98US-0020746.
PR 15-MAY-1997; 97US-0857216.
XX
PA (GETH) GENENTECH INC.
XX
PI Adams CM, Ashkenazi AJ, Chuntharapai A, Kim KJ;
XX WPI: 1999-045228/04.
DR N-PSDB: AAV72533.
XX
XX Human Apo-2 polypeptide inducing apoptosis - useful to treat
PT conditions linked with decreased apoptosis e.g. cancer, and produce
PT antibodies to increase or decrease apoptosis
XX
PS Example 14; Fig 16; 134pp; English.
XX
CC The present invention describes human Apo-2. Apo-2 can be used
CC therapeutically to induce apoptosis in mammalian cells, and so is useful

CC to treat conditions associated with decreased apoptosis e.g. cancer.
CC Apo-2 is believed to be a new tumour necrosis factor (TNF) receptor
CC (TNFR). TNF cytokines can induce apoptosis, thought to be initiated by
CC binding to TNFRs, and Apo-2 triggered caspase-dependent apoptosis. It
CC can be used to identify agents activating Apo-2, useful to treat
CC mammalian cancer cells, and to produce Apo-2 chimeras useful
CC therapeutically (e.g. those containing immunoglobulin sequences can be
CC inhibit apoptosis) or diagnostically (e.g. those comprising an epitope
CC tag polypeptide allow Apo-2 detection and purification using anti-tag
CC antibodies). It can be used to produce antibodies which can be combined
CC with a (particularly pharmaceutically acceptable) carrier in compositions
CC or used to produce dimeric molecules (especially homodimeric molecules
CC comprising first and second Apo-2 antibodies). Agonistic (especially
CC single-chain) antibodies can be administered to induce apoptosis in
CC mammalian cancer cells, and antagonistic antibodies used to block
CC excessive apoptosis (e.g. in neurodegenerative diseases). Apo-2
CC antibodies may also be used diagnostically e.g. to detect Apo-2
CC expression in cells/tissues and in Apo-2 purification. The present
CC sequence represents a single chain Apo-2 antibody, designated 20E6.
SO Sequence 312 AA;

Query Match 100.0%; Score 31; DB 20; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWMS 5
DB 70 sywms 74

RESULT 12
AAR85499
ID AAR85499 standard; Peptide: 5 AA.
XX
AC AAR85499;
XX
DT 16-MAR-1996 (first entry)
XX
DE ScFv(FWP51) CDR1H.
XX
XX Single chain antibody; scFv; antibody engineering; antitumour;
KW tumour antigen binding; cytotoxic T-lymphocyte; cell targeting;
KW monoclonal antibody; cancer; adoptive immunotherapy;
KW complementarity determining region; CDR.
XX
OS Synthetic.
XX
PN WO9530014-A1.
XX
PD 09-NOV-1995.
XX
PF 20-APR-1995; 95WO-EP01494.
XX
PR 02-MAY-1994; 94EP-0810244.
XX
PA (CIBA) CIBA GEIGY AG.
XX
PI Groner B, Moritz D;
XX WPI: 1995-393085/50.
DR
XX
XX New bifunctional proteins for use in killing tumour cells - contg. a
PT tumour antigen binding domain, a hinge region and a zeta chain
PT derived from a T-cell antigen receptor
XX
PS Disclosure; Page 6; 46pp; English.
XX
CC Complementarity determining regions (CDRs) CDR1H, CDR2H and CDR3H
CC (AAR85496-98) and CDR1L, CDR2L and CDR3L (AAR85502-04) are components
CC of the heavy chain variable region and light chain variable region,
CC respectively, of single chain antibody scFv(FWP51) (AAR85495). The CDRs

CC are derived from mouse monoclonal antibody FWP51 (ECACC 90112118)
CC directed to the human growth factor receptor HER2. The scfv is
CC used as the tumour antigen binding domain of a bifunctional protein
CC that is expressed in cytotoxic T-lymphocytes (CTLs), allowing
CC adoptive immunotherapy of cancer.

Sequence 5 AA:

Query Match 90.3%; Score 28; DB 16; Length 5;
Best Local Similarity 80.0%; Pred. No. 3.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
:||||
DB 1 tynms 5

RESULT 13

AAR97316
ID AAR97316 standard; peptide; 5 AA.

AC AAR97316;

DT 15-OCT-1996 (first entry)

DE Humanised monoclonal antibody heavy chain CDR.

KW Monoclonal antibody; humanised; mouse; framework region; FR; CDR;
KW complementary determining region; anti-carcinomaembryonic antigen;
KW CEA; diagnosis; imaging; therapy; immune response.

OS Mus musculus.

PN WO9611013-A1.

PD 18-APR-1996.

PF 28-SEP-1995; 95WO-US11964.

PR 05-OCT-1994; 94US-0318157.

PA (IMMU-) IMMUNOMEDICS INC.

PI Armour KL, Hansen HJ;

PS WPI; 1996-209653/21.

PT New humanised anti-CEA monoclonal antibody - having engrafted murine
PT CDRs, used for diagnosis, imaging and therapy of CEA-producing
PT cancers.

PS Claim 5; Page 38; 62pp; English.

CC New humanised monoclonal antibodies (MAbs) comprising the
CC complementary determining regions (CDRs) of a parental murine class
CC II anti-carcinomaembryonic (CEA) MAb engrafted to the framework
CC regions (FRs) of a heterologous antibody which can be derived from
CC any species including human, retain the anti-CEA binding specificity
CC of the parental murine MAb but are less immunogenic in a human
CC subject than the parental MAb. The humanised antibodies can be used
CC in diagnosis, imaging and therapy of CEA-producing cancers and
CC patients receiving the humanised antibodies and conjugates show
CC improved therapeutic results; decreased immune responses and
CC decreased immune-mediated adverse effects compared to the parent
CC antibody. This sequence is a murine derived CDR region which
CC corresponds to CDR 1 in the heavy chain of the humanised MAb. See
CC AAR97313-97333.

Sequence 5 AA:

Query Match 90.3%; Score 28; DB 17; Length 5;

Best Local Similarity 80.0%; Pred. No. 3.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
:||||
DB 1 tynms 5

RESULT 14

AA05038
ID AA05038 standard; peptide; 5 AA.

AC AA05038;

DT 16-JUN-1999 (first entry)

DE Tumour antigen antibody heavy chain CDR1 clone F15.

KW Tumour antigen; antibody; CDR; complementarity determining region;
KW binding molecule identification; tumour-specific binding polypeptide;
KW cancer therapy; heavy chain.

OS Homo sapiens.

PN WO9906834-A2.

PD 11-FEB-1999.

PF 04-AUG-1998; 98WO-US16280.

PR 04-AUG-1997; 97US-0905825.

PR 04-AUG-1997; 97US-1112222.

PA (IXSY-) IXSYS INC.

PI Huse WD, Watkins JD, Wu H;

PS WPI; 1999-153951/13.

DR N-PSDB; AAX28215.

PT Identifying binding molecules for ligands, particularly tumour
PT antigens - by selectively immobilising a population of binding
PT molecules to a solid support and screening for binding to two or
PT more ligands

PS Claim 15; Page 60; 80pp; English.

CC This sequence represents a heavy chain complementarity determining
CC region (CDR) from a tumour antigen specific antibody.
CC The invention relates to a method for identifying a binding molecule
CC having selective affinity for a ligand comprising: (a) selectively
CC immobilising a diverse population of binding molecules to a solid
CC support; (b) simultaneously contacting the diverse population immobilised
CC on the solid support with 2 or more ligands; and (c) determining at least
CC one binding molecule which selectively binds to one or more of the
CC ligands. The method allows for the rapid and efficient methods for the
CC identification of binding molecules which exhibit selective affinity for
CC one or more ligands of interest. They are used particularly for
CC identifying tumour-specific binding polypeptides which can be used as
CC targeting agents for cancer therapy that minimises impact on non-tumour
CC tissues.

Sequence 5 AA:

Query Match 90.3%; Score 28; DB 20; Length 5;
Best Local Similarity 80.0%; Pred. No. 3.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
:||||
DB 1 tynms 5

Search completed: November 19, 2001, 08:12:34
Job time: 87 sec

RESULT 15
AAV64781
ID AAV64781 standard; Protein: 59 AA.
XX
AC AAV64781;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human 5' EST related polypeptide SEQ ID NO:942.
XX
KW Human: 5' EST: expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification.
XX
OS Homo sapiens.
XX
PN M09953051-A2.
XX
PD 21-OCT-1999.
XX
PE 09-APR-1999; 99MO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR MPI; 2000-038446/03.
DR N-PSDB; AA242395.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 3: Page 641; 837pp; English.
XX
CC AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAV6451 to
CC AAV6438 represent the EST-related proteins corresponding to AA242265 to
CC AA243052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA242249 to AA242264 and AAV6444 to AAV6450 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 59 AA;

Query Match 90.3%; Score 28; DB 21; Length 59;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWMS 5
||||:
DB 50 symt 54

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:11:07 ; Search time 526.89 Seconds
(without alignments)
2.635 Million cell updates/sec

Title: US-09-610-118-61
Sequence: 1 SYWMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues
Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_MA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCrUS_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/2/paa/US060_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	5	1	PCT-US00-18152-61
2	31	100.0	5	20	US-09-610-118-61
3	31	100.0	5	22	US-09-832-312-61
4	31	100.0	27	22	US-60-160-209-3528
5	31	100.0	73	24	US-60-160-203-4861
6	31	100.0	98	1	PCT-US00-07946-664
7	31	100.0	98	1	PCT-US00-07946-666
8	31	100.0	98	5	US-08-124-469-32
9	31	100.0	98	8	US-08-472-771-32

10	31	100.0	98	8	US-08-486-202-32	Sequence 32, Appl
11	31	100.0	98	19	US-09-534-717-664	Sequence 664, App
12	31	100.0	98	19	US-09-534-717-666	Sequence 666, App
13	31	100.0	98	21	US-09-751-181-85	Sequence 85, Appl
14	31	100.0	116	21	US-09-758-455-598	Sequence 598, Appl
15	31	100.0	116	21	US-60-141-701-76	Sequence 76, Appl
16	31	100.0	123	24	US-60-141-701-77	Sequence 77, Appl
17	31	100.0	124	24	US-08-712-564-10	Sequence 10, Appl
18	31	100.0	125	24	US-60-141-701-75	Sequence 75, Appl
19	31	100.0	150	22	US-09-834-366-18232	Sequence 18232, A
20	31	100.0	150	24	US-60-197-873-18232	Sequence 1889, Ap
21	31	100.0	241	1	PCT-US01-19110-1889	Sequence 1889, Ap
22	31	100.0	241	22	US-09-880-748-1889	Sequence 1470, Ap
23	31	100.0	247	1	PCT-US01-19110-1470	Sequence 1470, Ap
24	31	100.0	247	22	US-09-880-748-1470	Sequence 1594, Ap
25	31	100.0	251	1	PCT-US01-19110-1594	Sequence 1594, Ap
26	31	100.0	251	22	US-09-880-748-1594	Sequence 10, Appl
27	31	100.0	312	14	US-09-079-029-10	Sequence 4698, Ap
28	31	100.0	335	21	US-09-738-626-4698	Sequence 3698, Ap
29	31	100.0	631	18	US-09-417-507-36498	Sequence 35861, A
30	31	100.0	789	18	PCT-US01-18510-4	Sequence 4, Appl1
31	28	90.3	5	1	US-08-318-157-23	Sequence 23, Appl
32	28	90.3	5	7	US-08-318-157A-23	Sequence 76, Appl
33	28	90.3	5	15	US-09-129-026-76	Sequence 23, Appl
34	28	90.3	5	16	US-09-253-794-23	Sequence 57, Appl
35	28	90.3	10	13	US-08-905-825-57	Sequence 6, Appl1
36	28	90.3	16	2	US-07-980-833-6	Sequence 48095, A
37	28	90.3	48	22	US-09-864-761-48095	Sequence 21288, A
38	28	90.3	46	24	US-60-236-359-21288	Sequence 478, App
39	28	90.3	59	14	US-09-069-047-478	Sequence 942, App
40	28	90.3	59	18	US-09-471-276-942	Sequence 23712, A
41	28	90.3	63	16	US-09-248-796-23712	Sequence 18292, A
42	28	90.3	69	22	US-09-834-366-18292	Sequence 18292, A
43	28	90.3	69	24	US-60-197-873-18292	Sequence 30906, A
44	28	90.3	72	18	US-09-417-507-30906	
45	28	90.3				

ALIGNMENTS

RESULT 1
PCT-US00-18152-61
Sequence 61, Appl
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-211-228
CURRENT APPLICATION NUMBER: PCT/US00/18152
CURRENT FILING DATE: 2000-06-30
EARLIER FILING DATE: 09/503,387
EARLIER FILING DATE: 2/14/00
EARLIER APPLICATION NUMBER: 09/454,824
EARLIER FILING DATE: 12/6/99
EARLIER APPLICATION NUMBER: 09/345,468
EARLIER FILING DATE: 6/30/99
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 61
LENGTH: 5
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-18152-61

Query Match 100.0%; Score 31; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYWMS 5
Db 1 SYWMS 5

```
RESULT 2
US-09-610-118-61
; Sequence 61, Application US/09610118
; GENERAL INFORMATION:
; APPLICANT: Bustfield, S.
; APPLICANT: Villaveal, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Valinchenker, W.
; APPLICANT: Gill, D.
; APPLICANT: Qian, M.
; APPLICANT: Kingsbury, G.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-211
; CURRENT APPLICATION NUMBER: US/09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2/14/00
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 12/6/99
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 6/30/99
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-610-118-61
```

```
Query Match          100.0%; Score 31; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYMS 5
Db 1 SYMS 5

RESULT 3
US-09-832-312-61
; Sequence 61, Application US/09832312
; GENERAL INFORMATION:
; APPLICANT: Bustfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-61
```

```
Query Match          100.0%; Score 31; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYMS 5
Db 1 SYMS 5
```

```
RESULT 4
US-60-160-209-3528
; Sequence 3528, Application US/60160209
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CL000113
; CURRENT APPLICATION NUMBER: US/60/160,209
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4646
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3528
; LENGTH: 27
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-209-3528
```

```
Query Match          100.0%; Score 31; DB 24; Length 27;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYMS 5
Db 17 SYMS 21
```

```
RESULT 5
US-60-160-203-4861
; Sequence 4861, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4861
; LENGTH: 73
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-4861
```

```
Query Match          100.0%; Score 31; DB 24; Length 73;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYMS 5
Db 53 SYMS 57
```

```
RESULT 6
PCT-US00-07946-664
; Sequence 664, Application PC/TUS0007946
; GENERAL INFORMATION:
; APPLICANT: Jochen, Salfeld et al.
; TITLE OF INVENTION: Human Antibodies That Bind Human IL-12 And Methods For Product
; FILE REFERENCE: BBI-0930PC
; CURRENT APPLICATION NUMBER: PCT/US00/07946
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/126,603
; EARLIER FILING DATE: March 25, 1999
; NUMBER OF SEQ ID NOS: 675
; SOFTWARE: PatentIn Ver. 2.0
```

SEQ ID NO 664
LENGTH: 98
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-07946-664

Query Match 100.0%; Score 31; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 7
PCT-US00-07946-666
Sequence 666, Application PC/TUS0007946
GENERAL INFORMATION:
APPLICANT: Jochen, Salfeld et al.
TITLE OF INVENTION: Human Antibodies That Bind Human IL-12 And Methods For Producing
FILE REFERENCE: BBI-093CPC
CURRENT APPLICATION NUMBER: PCT/US00/07946
CURRENT FILING DATE: 2000-03-24
EARLIER APPLICATION NUMBER: 60/126,603
EARLIER FILING DATE: March 25, 1999
NUMBER OF SEQ ID NOS: 675
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 666
LENGTH: 98
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-07946-666

Query Match 100.0%; Score 31; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 8
US-08-124-469-32
Sequence 32, Application US/08124469
GENERAL INFORMATION:
APPLICANT: RAPOPORT, Basil
APPLICANT: MCLACHLAN, Sandra M.
TITLE OF INVENTION: GRAVES' OPHTHALMOPATHY ASSOCIATED
TITLE OF INVENTION: ANTIBODIES, GRAVES' OPHTHALMOPATHY ORBITAL ANTIGEN,
TITLE OF INVENTION: AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR
STREET: 1455 PENNSYLVANIA AVE., N.W.
CITY: WASHINGTON,
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/124,469
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073

REFERENCE/DOCKET NUMBER: 102.105.301
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-124-469-32

Query Match 100.0%; Score 31; DB 5; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 9
US-08-472-771-32
Sequence 32, Application US/08472771
GENERAL INFORMATION:
APPLICANT: RAPOPORT, Basil
APPLICANT: MCLACHLAN, Sandra M.
TITLE OF INVENTION: GRAVES' OPHTHALMOPATHY ASSOCIATED
TITLE OF INVENTION: ANTIBODIES, GRAVES' OPHTHALMOPATHY ORBITAL ANTIGEN,
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR
STREET: 1455 PENNSYLVANIA AVE., N.W.
CITY: WASHINGTON,
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,771
FILING DATE: Herewith
PRIOR APPLICATION NUMBER:
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 102105.301DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-771-32

Query Match 100.0%; Score 31; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
|||||

Db 31 SYMS 35

RESULT 10

US-08-486-202-32
; Sequence 32, Application US/08486202
; GENERAL INFORMATION:
; APPLICANT: RAPOPORT, Basil
; APPLICANT: MCLACHLAN, Sandra M.
; TITLE OF INVENTION: GRAVES' OPHTHALMOPATHY ASSOCIATED
; TITLE OF INVENTION: ANTIBODIES, GRAVES' OPHTHALMOPATHY ORBITAL ANTIGEN,
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR
; STREET: 1455 PENNSYLVANIA AVE., N.W.
; CITY: WASHINGTON,
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,202
; FILING DATE: Herewith
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/124,469
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 102105,301DIY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8400
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-202-32

Query Match 100.0%; Score 31; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
Db 31 SYMS 35

RESULT 11

US-09-534-717-664
; Sequence 664, Application US/09534717
; GENERAL INFORMATION:
; APPLICANT: Jochen, Salfeld et al.
; TITLE OF INVENTION: Human Antibodies That Bind Human IL-12 And Methods For Producing
; FILE REFERENCE: Brl-093CP
; CURRENT APPLICATION NUMBER: US/09/534,717
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/126,603
; EARLIER FILING DATE: March 25, 1999
; NUMBER OF SEQ ID NOS: 675
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 664
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-534-717-664

Query Match 100.0%; Score 31; DB 19; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
Db 31 SYMS 35

RESULT 12

US-09-534-717-666
; Sequence 666, Application US/09534717
; GENERAL INFORMATION:
; APPLICANT: Jochen, Salfeld et al.
; TITLE OF INVENTION: Human Antibodies That Bind Human IL-12 And Methods For Producing
; FILE REFERENCE: Brl-093CP
; CURRENT APPLICATION NUMBER: US/09/534,717
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/126,603
; EARLIER FILING DATE: March 25, 1999
; NUMBER OF SEQ ID NOS: 675
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 666
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-534-717-666

Query Match 100.0%; Score 31; DB 19; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
Db 31 SYMS 35

RESULT 13

US-09-751-181-85
; Sequence 85, Application US/09751181
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp
; TITLE OF INVENTION: Specific Human Antibodies for Selective Cancer Therapy
; FILE REFERENCE: 456
; CURRENT APPLICATION NUMBER: US/09/751,181
; CURRENT FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 85
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-751-181-85

Query Match 100.0%; Score 31; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
Db 31 SYMS 35

RESULT 14

US-09-758-455-598
; Sequence 598, Application US/09758455
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

```

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM024
; CURRENT APPLICATION NUMBER: US/09/758,455
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 598
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-455-598

```

```

Query Match          100.0%; Score 31; DB 21; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYWMS 5
   |||||
Db 23 SYWMS 27

```

```

RESULT 15
US-60-141-701-76
; Sequence 76, Application US/60141701
; GENERAL INFORMATION:
; APPLICANT: Watkins, Brynmore
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO HIV-1 ENVELOPE
; FILE REFERENCE: 2026-4284
; CURRENT APPLICATION NUMBER: US/60/141,701
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-60-141-701-76

```

```

Query Match          100.0%; Score 31; DB 24; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYWMS 5
   |||||
Db 32 SYWMS 36

```

Search completed: November 19, 2001, 08:23:14
 Job time: 727 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:11:08 ; Search time 11.44 Seconds
(without alignments)
2.328 Million cell updates/sec

Title: US-09-610-118-61

Perfect score: 31
Sequence: 1 SYWMS 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 25229 seqs, 5326477 residues

Total number of hits satisfying chosen parameters: 25229

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_US01_32234-11
2: /cgn2_6/ptodata/1/paa/US06_NEM_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEM_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEM_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEM_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US060_NEM_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	90.3	472	5	US-09-897-516-7279
2	27	87.1	107	1	PCT-US01-32234-11
3	27	87.1	181	5	US-09-897-516-6974
4	27	87.1	463	4	US-08-721-612C-3
5	27	87.1	464	4	US-08-721-612C-9
6	27	87.1	470	5	US-09-815-242-12078
7	27	87.1	1346	5	US-09-657-440-4
8	25	80.6	253	5	US-09-786-039-4
9	25	80.6	255	5	US-09-786-039-5
10	25	80.6	324	6	US-60-325-537-24
11	25	80.6	478	5	US-09-815-242-5719
12	25	80.6	478	5	US-09-815-242-12480
13	25	80.6	497	5	US-09-897-516-5684
14	25	80.6	2161	5	US-09-574-141A-3
15	25	80.6	2329	5	US-09-897-516-4693
16	24	77.4	7	5	US-09-560-390-28
17	24	77.4	54	5	US-09-897-516-4882
18	24	77.4	97	5	US-09-897-516-6863
19	24	77.4	98	1	PCT-US01-09369-24
20	24	77.4	116	5	US-09-741-843-4
21	24	77.4	116	5	US-09-741-843-8
22	24	77.4	116	5	US-09-741-843-9
23	24	77.4	134	5	US-09-897-516-5475
24	24	77.4	172	5	US-09-897-516-7377
25	24	77.4	255	5	US-09-978-197-2
26	24	77.4	258	5	US-09-897-516-6309
27	24	77.4	265	5	US-09-978-197-4

28	24	77.4	275	5	US-09-803-671B-5	Sequence 5, Appl1
29	24	77.4	277	5	US-09-803-671B-6	Sequence 6, Appl1
30	24	77.4	277	5	US-09-803-671B-7	Sequence 7, Appl1
31	24	77.4	282	5	US-09-897-516-5639	Sequence 5639, Ap
32	24	77.4	283	5	US-09-897-516-7996	Sequence 7996, Ap
33	24	77.4	295	5	US-09-815-242-5269	Sequence 5269, Ap
34	24	77.4	298	5	US-09-815-242-5385	Sequence 5385, Ap
35	24	77.4	299	5	US-09-815-242-12197	Sequence 12197, A
36	24	77.4	299	5	US-09-815-242-12958	Sequence 12958, A
37	24	77.4	310	5	US-09-815-242-13088	Sequence 13088, A
38	24	77.4	310	5	US-09-815-242-10910	Sequence 10910, A
39	24	77.4	317	5	US-09-825-561A-84	Sequence 5631, Ap
40	24	77.4	329	5	US-09-897-516-5090	Sequence 5090, Ap
41	24	77.4	354	5	US-09-966-521-12	Sequence 12, Appl
42	24	77.4	359	5	US-09-966-521-12	Sequence 6980, Ap
43	24	77.4	368	5	US-09-897-516-6980	Sequence 78, Appl
44	24	77.4	370	5	US-09-966-521-78	Sequence 11724, A
45	24	77.4	395	5	US-09-815-242-11724	

ALIGNMENTS

```
RESULT 1
US-09-897-516-7279
; Sequence 7279, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Cordin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Salter, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 7279
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.

Query Match      90.3%  Score 28:  DB 5:  Length 472:
Best Local Similarity 80.0%:  Pred. No. 62:
Matches 4:  Conservative 1:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 SYWMS 5
Db 105 TYWMS 109

RESULT 2
PCT-US01-32234-11
; Sequence 11, Application PC/TUS0132234
; GENERAL INFORMATION:
; APPLICANT: von Bdingen, Hans-Christian
; APPLICANT: Genath, Claude P.
; APPLICANT: Hauser, Stephen L.
; TITLE OF INVENTION: Recombinant Antibody Fragments as Autoantibody
; FILE REFERENCE: SF01-025-2
; CURRENT APPLICATION NUMBER: PCT/US01/32234
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/691,654
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
```

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 107
TYPE: PRT
ORGANISM: human
PCT-US01-32234-11

Query Match 87.1% Score 27; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5
DB 32 YWMS 35

RESULT 3
US-09-897-516-6974
Sequence 6974, Application US/09897516
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Goldmen, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spilidonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 6974
LENGTH: 181
TYPE: PRT
ORGANISM: Xenorhabdus sp.
US-09-897-516-6974

Query Match 87.1% Score 27; DB 5; Length 181;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWMS 5
DB 135 YWMS 138

RESULT 4
US-08-721-612C-3
Sequence 3, Application US/08721612C
GENERAL INFORMATION:
APPLICANT: Mueller, John P.
Evans, Mark J.
Mueller, Eileen Elliott
Rollins, Scott
Rother, Russell P.
Matis, Louis A.
TITLE OF INVENTION: Porcine Cell Interaction Proteins
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel, Alexion Pharmaceuticals
STREET: 25 Science Park, Suite 360
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4mb storage

COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 95
SOFTWARE: MICROSOFT WORD 97 SR-2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,612C
FILING DATE: 26-Sep-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/004489
FILING DATE: 28-SEPTEMBER-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A...
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-154
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: protein
DESCRIPTION: 2A2 Heavy Chain "cDNA" product
US-08-721-612C-3

Query Match 87.1% Score 27; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWM 4
DB 50 SYWM 53

RESULT 5
US-08-721-612C-9
Sequence 9, Application US/08721612C
GENERAL INFORMATION:
APPLICANT: Mueller, John P.
Evans, Mark J.
Mueller, Eileen Elliott
Rollins, Scott
Rother, Russell P.
Matis, Louis A.
TITLE OF INVENTION: Porcine Cell Interaction Proteins
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel, Alexion Pharmaceuticals
STREET: 25 Science Park, Suite 360
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4mb storage
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 95
SOFTWARE: MICROSOFT WORD 97 SR-2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,612C
FILING DATE: 26-Sep-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/004489
FILING DATE: 28-SEPTEMBER-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A...
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-154
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790

TELEFAX: (203) 772-3655
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 464 amino acids
 TYPE: Amino acid
 STRANDEDNESS: Single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: 3A4 heavy chain "cDNA" product
 SEQUENCE DESCRIPTION: SEQ ID NO: 9:
 US-08-721-612C-9

Query Match 87.1%; Score 27; DB 4; Length 464;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SYMM 4
 1111
 Db 50 SYMM 53

RESULT 6
 US-09-815-242-12078
 : Sequence 12078, Application US/09815242
 : GENERAL INFORMATION:
 : APPLICANT: Haselbeck, Robert
 : APPLICANT: Ohlsen, Karl L.
 : APPLICANT: Zyskind, Judith W.
 : APPLICANT: Wall, Daniel
 : APPLICANT: Trawick, John D.
 : APPLICANT: Carr, Grant J.
 : APPLICANT: Yamamoto, Robert T.
 : APPLICANT: Xu, H. Howard
 : TITLE OF INVENTION: Identification of Essential Genes in
 : FILE REFERENCE: ELITRA.011A
 : CURRENT APPLICATION NUMBER: US/09/815,242
 : PRIOR FILING DATE: 2001-03-21
 : PRIOR APPLICATION NUMBER: 60/191,078
 : PRIOR FILING DATE: 2000-03-21
 : PRIOR APPLICATION NUMBER: 60/206,848
 : PRIOR FILING DATE: 2000-05-23
 : PRIOR APPLICATION NUMBER: 60/207,727
 : PRIOR FILING DATE: 2000-05-26
 : PRIOR APPLICATION NUMBER: 60/242,578
 : PRIOR FILING DATE: 2000-10-23
 : PRIOR APPLICATION NUMBER: 60/253,625
 : PRIOR FILING DATE: 2000-11-27
 : PRIOR APPLICATION NUMBER: 60/257,931
 : PRIOR FILING DATE: 2000-12-22
 : PRIOR APPLICATION NUMBER: 60/269,308
 : PRIOR FILING DATE: 2001-02-16
 : NUMBER OF SEQ ID NOS: 14110
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 12078
 : LENGTH: 470
 : TYPE: PRT
 : ORGANISM: Pseudomonas aeruginosa
 : US-09-815-242-12078

Query Match 87.1%; Score 27; DB 5; Length 470;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SYMM 5
 1111
 Db 99 SYMM 102

RESULT 7
 US-09-657-440-4

: Sequence 4, Application US/09657440
 : GENERAL INFORMATION:
 : APPLICANT: ASHLEY, Gary
 : APPLICANT: BETLACH, Melanie C.
 : APPLICANT: BETLACH, Mary C.
 : APPLICANT: MCDANIEL, Robert
 : APPLICANT: TANG, Li
 : TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
 : FILE REFERENCE: 300622002120
 : CURRENT APPLICATION NUMBER: US/09/657,440
 : CURRENT FILING DATE: 2000-09-07
 : PRIOR APPLICATION NUMBER: 09/320,878
 : PRIOR FILING DATE: 1999-05-27
 : PRIOR APPLICATION NUMBER: CIP OF 09/141,908
 : PRIOR FILING DATE: 1998-08-28
 : NUMBER OF SEQ ID NOS: 34
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 4
 : LENGTH: 1346
 : TYPE: PRT
 : ORGANISM: Streptomyces venezuelae
 : US-09-657-440-4

Query Match 87.1%; Score 27; DB 5; Length 1346;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMM 5
 1111
 Db 902 SYMM 906

RESULT 8
 US-09-786-039-4
 : Sequence 4, Application US/09786039
 : GENERAL INFORMATION:
 : APPLICANT: Brophy, Bridgid
 : APPLICANT: Langley, Brett
 : APPLICANT: L'Huillier, Phillip John
 : TITLE OF INVENTION: Mammary-Specific Transcription Factor
 : FILE REFERENCE: 0208/01254-US0
 : CURRENT APPLICATION NUMBER: US/09/786,039
 : CURRENT FILING DATE: 1999-08-30
 : PRIOR APPLICATION NUMBER: NZ 331612
 : PRIOR FILING DATE: 1998-08-28
 : NUMBER OF SEQ ID NOS: 25
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 4
 : LENGTH: 253
 : TYPE: PRT
 : ORGANISM: Murine
 : US-09-786-039-4

Query Match 80.6%; Score 25; DB 5; Length 253;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYMM 5
 1111
 Db 44 SYMM 48

RESULT 9
 US-09-786-039-5
 : Sequence 5, Application US/09786039
 : GENERAL INFORMATION:
 : APPLICANT: Brophy, Bridgid
 : APPLICANT: Langley, Brett
 : APPLICANT: L'Huillier, Phillip John
 : TITLE OF INVENTION: Mammary-Specific Transcription Factor

```
; TITLE OF INVENTION: Bfets
; FILE REFERENCE: 0208/01254-US0
; CURRENT APPLICATION NUMBER: US/09/786,039
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: NZ 331612
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Bovine
US-09-786-039-5
```

```
Query Match
Best Local Similarity 80.6%; Score 25; DB 5; Length 255;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 SYMS 5
Db 44 SYMS 48
```

```
RESULT 10
US-60-325-537-24
; Sequence 24, Application US/60325537
; GENERAL INFORMATION:
; APPLICANT: Stein, Joshua C
; APPLICANT: Cso, Yongwei
; TITLE OF INVENTION: Methods for Analyzing Biological Elements
; FILE REFERENCE: 16517,246/38-10(52045)A
; CURRENT APPLICATION NUMBER: US/60/325,537
; CURRENT FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-60-325-537-24
```

```
Query Match
Best Local Similarity 80.6%; Score 25; DB 6; Length 324;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYMS 5
Db 240 SYMS 244
```

```
RESULT 11
US-09-815-242-5719
; Sequence 5719, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5719
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5719
```

```
Query Match
Best Local Similarity 80.6%; Score 25; DB 5; Length 478;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYMS 5
Db 169 SYMS 173
```

```
RESULT 12
US-09-815-242-12480
; Sequence 12480, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12480
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12480
```

```
Query Match
Best Local Similarity 80.6%; Score 25; DB 5; Length 478;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYMS 5
```

Db 169 TWM 173

```

RESULT 13
US-09-897-516-5684
; Sequence 5684, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spilldonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 5684
; LENGTH: 497
; TYPE: PRF
; ORGANISM: Xenorhabdus sp.
US-09-897-516-5684

```

```

Query Match      80.6%; Score 25; DB 5; Length 497;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SYWMS 5
Db 425 SYWMS 429

```

```

RESULT 14
US-09-574-141A-3
; Sequence 3, Application US/09574141A
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Meng, Baozhong
; TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS
; FILE REFERENCE: 07678/035005
; CURRENT APPLICATION NUMBER: US/09/574,141A
; CURRENT FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 60/047,147
; PRIOR FILING DATE: 1997-05-20
; PRIOR APPLICATION NUMBER: 60/069,902
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 09/081,320
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2161
; TYPE: PRF
; ORGANISM: Rupestris stem pitting associated virus
US-09-574-141A-3

```

```

Query Match      80.6%; Score 25; DB 5; Length 2161;
Best Local Similarity 60.0%; Pred. No. 7.1e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWMS 5
Db 482 SYWMS 486

```

```

RESULT 15
US-09-897-516-4693
; Sequence 4693, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spilldonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 4693
; LENGTH: 2329
; TYPE: PRF
; ORGANISM: Xenorhabdus sp.
US-09-897-516-4693

```

```

Query Match      80.6%; Score 25; DB 5; Length 2329;
Best Local Similarity 60.0%; Pred. No. 7.6e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYWMS 5
Db 1449 SYWMS 1453

```

Search completed: November 19, 2001, 08:23:31
Job time: 743 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:11:07 : Search time 43.5 Seconds
(without alignments)
2.587 Million cell updates/sec

Title: US-09-610-118-61

Perfect score: 31

Sequence: 1 SYWMS 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	98	1	US-07-942-245-37
2	31	100.0	117	3	US-08-545-809A-95
3	28	90.3	5	2	US-08-318-157B-23
4	28	90.3	10	1	US-08-208-886C-83
5	28	90.3	10	1	US-08-208-886C-87
6	28	90.3	10	1	US-08-704-744-85
7	28	90.3	10	1	US-08-704-744-89
8	28	90.3	10	1	US-08-469-557-64
9	28	90.3	10	1	US-08-469-557-68
10	28	90.3	10	2	US-08-290-793B-64
11	28	90.3	10	2	US-08-290-793B-68
12	28	90.3	118	3	US-08-767-128-2
13	28	90.3	118	3	US-08-767-128-8
14	28	90.3	119	2	US-08-318-157B-2
15	28	90.3	119	2	US-08-318-157B-8
16	28	90.3	119	2	US-08-318-157B-9
17	28	90.3	119	2	US-08-318-157B-10
18	28	90.3	119	2	US-08-318-157B-11
19	28	90.3	119	2	US-08-318-157B-12
20	28	90.3	119	2	US-08-318-157B-13
21	28	90.3	119	2	US-08-318-157B-14
22	28	90.3	119	2	US-08-318-157B-15
23	28	90.3	119	2	US-08-318-157B-17
24	28	90.3	119	2	US-08-318-157B-57
25	28	90.3	119	2	US-08-318-157B-58
26	28	90.3	119	3	US-08-767-128-6
27	28	90.3	218	2	US-08-399-889-25

28	28	90.3	218	3	US-09-167-364-25	Sequence 25, Appl
29	28	90.3	218	4	US-09-439-897-4	Sequence 4, Appl
30	28	90.3	241	1	US-08-235-838-11	Sequence 11, Appl
31	28	90.3	241	2	US-08-465-473B-11	Sequence 11, Appl
32	28	90.3	298	4	US-09-318-661-2	Sequence 2, Appl
33	28	90.3	467	1	US-08-704-744-81	Sequence 81, Appl
34	28	90.3	471	2	US-08-399-889-24	Sequence 24, Appl
35	28	90.3	471	4	US-09-167-364-24	Sequence 24, Appl
36	28	90.3	503	3	US-08-700-651-6	Sequence 6, Appl
37	28	90.3	637	1	US-08-235-838-16	Sequence 16, Appl
38	28	90.3	637	2	US-08-465-473B-16	Sequence 5, Appl
39	28	90.3	636	1	US-08-765-081-5	Sequence 5, Appl
40	28	90.3	636	3	US-09-098-082-5	Sequence 5, Appl
41	28	90.3	696	5	PCT-US95-06994-5	Sequence 5, Appl
42	28	90.3	703	5	PCT-US95-06994-8	Sequence 6, Appl
43	28	90.3	718	5	PCT-US95-06994-6	Sequence 6, Appl
44	28	90.3	1694	1	US-08-494-168-2	Sequence 2, Appl
45	28	90.3				

ALIGNMENTS

RESULT 1
US-07-942-245-37
Sequence 37, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: BRAYDON C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrie, Mion, Zimm, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3302
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-37

Query Match 100.0%; Score 31; DB 1; Length 98;

Best Local Similarity 100.0%; Pred. No. 61;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYWMS 5
|||||
DB 31 SYWMS 35

RESULT 2
US-08-545-809A-95
; Sequence 95, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Matsuda, Fumihiko
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,809A
; FILING DATE: 27-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP93/00603
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 06501/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-545-809A-95

Query Match 100.0%; Score 31; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
DB 50 SYWMS 54

RESULT 3
US-08-318-157B-23
; Sequence 23, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-157B-23

Query Match 90.3%; Score 28; DB 2; Length 5;
Best Local Similarity 80.0%; Pred. No. 1.6e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
DB 1 SYWMS 5

RESULT 4
US-08-208-886C-83
; Sequence 83, Application US/08208886C
; Patent No. 5597710
; GENERAL INFORMATION:
; APPLICANT: Dalle, Barbara
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Tindall, Stephen
; TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,886C
; FILING DATE: March 10, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Foulke, Cynthia L.
; REGISTRATION NUMBER: 32,364
; REFERENCE/DOCKET NUMBER: JB0429
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908 298 2987
; TELEFAX: 908 298 5388
; INFORMATION FOR SEQ ID NO: 83:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-208-886C-83

Query Match 90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
||||:
DB 6 SYMT 10

RESULT 5
US-08-208-886C-87
Sequence 87, Application US/08208886C
Patent No. 5597710
GENERAL INFORMATION:
APPLICANT: Dalle, Barbara
APPLICANT: Miller, Kenneth
APPLICANT: Murgolo, Nicholas
APPLICANT: Tindall, Stephen
TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/208-886C
FILING DATE: March 10, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: JB0429
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908 298 2987
TELEFAX: 908 298 5388
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-208-886C-87

Query Match 90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
||||:
DB 6 SYMT 10

RESULT 6

US-08-704-744-85
Sequence 85, Application US/08704744
Patent No. 5705154

GENERAL INFORMATION:
APPLICANT: Dalle, Barbara
APPLICANT: Miller, Kenneth
APPLICANT: Murgolo, Nicholas
APPLICANT: Tindall, Stephen
TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.3
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/704,744
FILING DATE: 06-SEPT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208886
FILING DATE: 10-MAR-1994
APPLICATION NUMBER: PCT/US/95/02400
FILING DATE: 08-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: JB0429K
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 298-2987
TELEFAX: (908) 298-5388
TELEX:
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-704-744-85

Query Match 90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYMS 5
||||:
DB 6 SYMT 10

RESULT 7
US-08-704-744-89
Sequence 89, Application US/08704744
Patent No. 5705154
GENERAL INFORMATION:
APPLICANT: Dalle, Barbara
APPLICANT: Miller, Kenneth
APPLICANT: Murgolo, Nicholas
APPLICANT: Tindall, Stephen
TITLE OF INVENTION: Humanized Monoclonal Antibodies Against Human Interleukin-4
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey

COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.3
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/704,744
FILING DATE: 06-SEPT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208886
FILING DATE: 10-MAR-1994
APPLICATION NUMBER: PCT/US/95/02400
FILING DATE: 08-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: JB0429K
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 298-2987
TELEFAX: (908) 298-5388
TELEX:
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-704-744-89

Query Match 90.3% Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 6 SYWMT 10

RESULT 8
US-08-469-557-64
Sequence 64, Application US/08469557
Patent No. 5770403
GENERAL INFORMATION:
APPLICANT: Dalle, Barbara
APPLICANT: Le, Hung
APPLICANT: Miller, Kenneth
APPLICANT: Murgolo, Nicholas
APPLICANT: Nguyen, Hanh
APPLICANT: Tindall, Stephen
TITLE OF INVENTION: Cloning and Expression of
TITLE OF INVENTION: Humanized Monoclonal Antibodies
TITLE OF INVENTION: Against Human Interleukin-4
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,557
FILING DATE: 06-JUN-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,793
FILING DATE: August 16, 1994
APPLICATION NUMBER: PCT/US93/01301
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/841,659
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/782,784
FILING DATE: 24-OCT-1991
APPLICATION NUMBER: US 07/499,327
FILING DATE: 21-MAY-1990
APPLICATION NUMBER: PCT/US88/03631
FILING DATE: 21-OCT-1988
APPLICATION NUMBER: US 07/655,966
FILING DATE: 14-FEB-1991
APPLICATION NUMBER: US 07/113,623
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: US 06/881,553
FILING DATE: 03-JUL-1986
APPLICATION NUMBER: US 06/843,958
FILING DATE: 25-MAR-1986
APPLICATION NUMBER: US 06/799,668
FILING DATE: 19-NOV-1985
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: 2409K7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908 298-2987
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-557-64

Query Match 90.3% Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 6 SYWMT 10

RESULT 9
US-08-469-557-68
Sequence 68, Application US/08469557
Patent No. 5770403
GENERAL INFORMATION:
APPLICANT: Dalle, Barbara
APPLICANT: Le, Hung
APPLICANT: Miller, Kenneth
APPLICANT: Murgolo, Nicholas
APPLICANT: Nguyen, Hanh
APPLICANT: Tindall, Stephen
TITLE OF INVENTION: Cloning and Expression of
TITLE OF INVENTION: Humanized Monoclonal Antibodies
TITLE OF INVENTION: Against Human Interleukin-4
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,557
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,793
FILING DATE: August 16, 1994
APPLICATION NUMBER: PCT/US93/01301
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/841,659
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/782,784
FILING DATE: 24-OCT-1991
APPLICATION NUMBER: US 07/499,327
FILING DATE: 21-MAY-1990
APPLICATION NUMBER: PCT/US88/03631
FILING DATE: 21-OCT-1988
APPLICATION NUMBER: US 07/655,966
FILING DATE: 14-FEB-1991
APPLICATION NUMBER: US 07/113,623
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: US 06/881,553
FILING DATE: 03-JUL-1986
APPLICATION NUMBER: US 06/843,958
FILING DATE: 25-MAR-1986
APPLICATION NUMBER: US 06/799,668
FILING DATE: 19-NOV-1985
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: 2409K7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908 298-2987
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-557-68

Query Match 90.3%; Score 28; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
DB 6 SYWMT 10

RESULT 10
US-08-290-793B-64
; Sequence 64, Application US/08290793B
; Patent No. 5863537
; GENERAL INFORMATION:
; APPLICANT: Dalie, Barbara
; APPLICANT: Le, Hung
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Nguyen, Hanh
; APPLICANT: Tindall, Stephen
; APPLICANT: Zavodny, Paul
; TITLE OF INVENTION: Cloning and Expression of
; TITLE OF INVENTION: Humanized Monoclonal Antibodies
; TITLE OF INVENTION: Against Human Interleukin-4
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,793B
FILING DATE: August 16, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01301
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/841,659
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/782,784
FILING DATE: 24-OCT-1991
APPLICATION NUMBER: US 07/499,327
FILING DATE: 21-MAY-1990
APPLICATION NUMBER: PCT/US88/03631
FILING DATE: 21-OCT-1988
APPLICATION NUMBER: US 07/655,966
FILING DATE: 14-FEB-1991
APPLICATION NUMBER: US 07/113,623
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: US 06/881,553
FILING DATE: 03-JUL-1986
APPLICATION NUMBER: US 06/843,958
FILING DATE: 25-MAR-1986
APPLICATION NUMBER: US 06/799,668
FILING DATE: 19-NOV-1985
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: 2409K7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908 298-2987
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-290-793B-64

Query Match 90.3%; Score 28; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
DB 6 SYWMT 10

RESULT 11
US-08-290-793B-68
; Sequence 68, Application US/08290793B
; Patent No. 5863537
; GENERAL INFORMATION:
; APPLICANT: Dalie, Barbara
; APPLICANT: Le, Hung
; APPLICANT: Miller, Kenneth
; APPLICANT: Murgolo, Nicholas
; APPLICANT: Nguyen, Hanh
; APPLICANT: Tindall, Stephen
; APPLICANT: Zavodny, Paul
; TITLE OF INVENTION: Cloning and Expression of

;; TITLE OF INVENTION: Humanized Monoclonal Antibodies
;; TITLE OF INVENTION: Against Human Interleukin-4
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corporation
;; STREET: 2000 Gallopington Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07033-0530
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: Macintosh 6.0.5
;; SOFTWARE: Microsoft Word 5.1a
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/290,793B
;; FILING DATE: August 16, 1994
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/01301
;; FILING DATE: 19-FEB-1992
;; APPLICATION NUMBER: US 07/841,659
;; FILING DATE: 19-FEB-1992
;; APPLICATION NUMBER: US 07/782,784
;; FILING DATE: 24-OCT-1991
;; APPLICATION NUMBER: US 07/499,327
;; FILING DATE: 21-MAY-1990
;; APPLICATION NUMBER: PCT/US88/03631
;; FILING DATE: 21-OCT-1988
;; APPLICATION NUMBER: US 07/655,966
;; FILING DATE: 14-FEB-1991
;; APPLICATION NUMBER: US 07/113,623
;; FILING DATE: 26-OCT-1987
;; APPLICATION NUMBER: US 06/881,553
;; FILING DATE: 03-JUL-1986
;; APPLICATION NUMBER: US 06/843,958
;; FILING DATE: 25-MAR-1986
;; APPLICATION NUMBER: US 06/799,668
;; FILING DATE: 19-NOV-1985
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Foulke, Cynthia L.
;; REGISTRATION NUMBER: 32,364
;; REFERENCE/DOCKET NUMBER: 2409K7
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908 298-2387
;; TELEFAX: 908-298-5388
;;
;; INFORMATION FOR SEQ ID NO: 68:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;;
;; US-08-290-793B-68

Query Match 90.3%; Score 28; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
|||||
DB 6 SYMWT 10

RESULT 12
US-08-767-128-2
; Sequence 2, Application US/08767128
; Patent No. 6111079
; GENERAL INFORMATION:
; APPLICANT: WYLLIE, DWANE E.
; APPLICANT: LOPEZ, OSVALDO
; APPLICANT: MURRAY, PETER JOSEPH
; APPLICANT: GOEBEL, PETER

;; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND
;; TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE
;; NUMBER OF SEQUENCES: 46
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
;; STREET: 3100 No. 6111079west Center, 90 South Seventh St
;; CITY: Minneapolis
;; STATE: MN
;; COUNTRY: USA
;; ZIP: 55402
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: IBM Compatible
;; SOFTWARE: FastSeq Version 1.5
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/767,128
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE: 04-DEC-1996
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US96/09258
;; FILING DATE: 05-JUN-1996
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/541,373
;; FILING DATE: 10-OCT-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/462,798
;; FILING DATE: 05-JUN-1995
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Carter, Charles G.
;; REGISTRATION NUMBER: 35,093
;; REFERENCE/DOCKET NUMBER: 8648.49USF1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 612/371-5278
;; TELEFAX: 612/332-9081
;;
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 118 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: Internal
;; ORIGINAL SOURCE:
;;
;; US-08-767-128-2

Query Match 90.3%; Score 28; DB 3; Length 118;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYMS 5
|||||
DB 31 SYMWT 35

RESULT 13
US-08-767-128-8
; Sequence 8, Application US/08767128
; Patent No. 6111079
; GENERAL INFORMATION:
; APPLICANT: WYLLIE, DWANE E.
; APPLICANT: LOPEZ, OSVALDO
; APPLICANT: MURRAY, PETER JOSEPH
; APPLICANT: GOEBEL, PETER
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND

TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 6111079west Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,128
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 04-DEC-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09258
FILING DATE: 05-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/541,373
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,798
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648,49USF1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/371-5278
TELEFAX: 612/332-9081
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-767-128-8

Query Match 90.3%; Score 28; DB 3; Length 118;
Best Local Similarity 80.0%; Pred No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 31 SYWMN 35

RESULT 14
US-08-318-157B-2
Sequence 2, Application US/08318157B
Patent No. 5874540
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
APPLICANT: ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
TITLE OF INVENTION: MOOSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-2

Query Match 90.3%; Score 28; DB 2; Length 119;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
|||||
DB 31 SYWMN 35

RESULT 15
US-08-318-157B-8
Sequence 8, Application US/08318157B
Patent No. 5874540
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
APPLICANT: ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399
 TELE: 904136
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 119 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-318-157B-8

Query Match 90.3%; Score 28; DB 2; Length 119;
 Best Local Similarity 80.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYWMS 5
 :|||||
 Db 31 TYWMS-35

Search completed: November 19, 2001, 08:13:24
 Job time: 137 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:14:20 : Search time 48.99 Seconds
(without alignments)
26.433 Million cell updates/sec

Title: US-09-610-118-62

Perfect score: 89

Sequence: 1 NIKODGSEKYYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	96.6	147	2	I37780
2	82	92.1	97	2	PH0872
3	82	92.1	98	2	PL0123
4	82	92.1	110	2	PH1655
5	82	92.1	114	2	S36280
6	82	92.1	117	2	S78486
7	82	92.1	117	2	S17079
8	82	92.1	122	2	S31675
9	82	92.1	136	2	S31587
10	82	92.1	139	2	I37781
11	79	88.8	140	2	S22657
12	77	86.5	123	2	S31509
13	67	75.3	136	2	S16847
14	67	75.3	136	2	S60296
15	66	74.2	98	2	S29543
16	66	74.2	117	2	S36259
17	66	74.2	134	2	S31688
18	66	74.2	135	2	S31598
19	66	74.2	140	2	S70442
20	65	73.0	121	2	PH1661
21	64	71.9	94	2	PL0120
22	64	71.9	97	2	S44115
23	64	71.9	98	2	PL0116
24	64	71.9	98	2	S29546
25	64	71.9	109	2	PH1644
26	64	71.9	109	2	PH1646
27	64	71.9	111	2	PH1643
28	64	71.9	111	2	PH1645
29	64	71.9	113	2	S38490

30	64	71.9	114	2	S46390	Ig heavy chain V r
31	64	71.9	114	2	S46392	Ig heavy chain V r
32	64	71.9	117	2	S36270	Ig heavy chain V r
33	64	71.9	118	2	S31677	Ig heavy chain V r
34	64	71.9	118	2	S31116	Ig heavy chain - h
35	64	71.9	119	2	S36005	Ig heavy chain V r
36	64	71.9	120	2	S31112	Ig heavy chain - h
37	64	71.9	121	2	S19666	Ig heavy chain V r
38	64	71.9	121	2	S36005	Ig heavy chain V r
39	64	71.9	122	2	S36005	Ig heavy chain V r
40	64	71.9	122	2	S38493	Ig heavy chain - h
41	64	71.9	123	2	S31119	Ig heavy chain - h
42	64	71.9	130	2	PL0098	Ig heavy chain pre
43	64	71.9	132	2	S31603	Ig heavy chain V r
44	64	71.9	134	2	S31679	Ig heavy chain V r
45	64	71.9	139	2	S31674	Ig heavy chain V r

ALIGNMENTS

RESULT 1
I37780
Ig variable region (VDJ) (clone T20-11) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
C:Accession: I37780; S25474
R:Demaison, C.; Chastagner, P.; Theze, J.; Zouali, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
A:Title: Somatic diversification in the heavy chain variable region genes expressed b
A:Reference number: A36876; MUID:94119917
A:Accession: I37780
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-147 <RES>
A:Cross-references: EMBL:X67943; NID:933578; PIDD:CAA48130.1; PID:933579
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
F:28-111/Domain: Immunoglobulin homology <IMM>

Query Match 96.6%; Score 86; DB 2; Length 147;
Best Local Similarity 94.1%; Pred. No. 3.8e-07;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIKODGSEKYYADSVRG 17
Db 63 NIKODGSEKYYADSVRG 79

RESULT 2
PH0872
Ig heavy chain V region (anti-DNA, III-3R) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 16-Aug-1996
C:Accession: PH0872
R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiot
A:Reference number: PH0862; MUID:92078875
A:Accession: PH0872
A:Molecule type: DNA
A:Residues: 1-97 <MAN>
A:Note: a "G" was inserted at the position of between 72nd and 73rd of the DNA sequen
C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that be
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>
F:30-35/Region: complementarity-determining 1
F:49-66/Region: complementarity-determining 2

Query Match 92.1%; Score 82; DB 2; Length 97;
Best Local Similarity 88.2%; Pred. No. 1.1e-06;

```
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKODSEKRYVDSVKG 17
   |||||
Db 49 NIKODSEKRYVDSVKG 65

RESULT 3
PI0123
Ig heavy chain V-III region (TD-Vr) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
C:Accession: PI0123; S26897
R:Bird, J.; Gallili, N.; Link, M.; Stiles, D.; Sklar, J.
J. Exp. Med. 168, 223-245, 1988
A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin
A:Reference number: PI0116; MUID:88286083
A:Accession: PI0123
A:Molecule type: mRNA
A:Residues: 1-98 <BIR>
A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26897
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:212354; NID:932930; PIDD:CA78224.1; PID:932931
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:49-65/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match
Best Local Similarity 92.1%; Score 82; DB 2; Length 98;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKODSEKRYVDSVKG 17
   |||||
Db 50 NIKODSEKRYVDSVKG 66

RESULT 4
PH1655
Ig heavy chain V region (clone 2E8) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C:Accession: PH1655
R:Hillson, J.L.; Kart, N.S.; Oppliger, I.R.; Mannik, M.; Sasso, E.H.
J. Exp. Med. 178, 331-336, 1993
A:Title: The structural basis of germ-line-encoded VH3 immunoglobulin binding to staphylo
A:Reference number: PH1642; MUID:93301610
A:Accession: PH1655
A:Molecule type: mRNA
A:Residues: 1-110 <HIL>
A:Experimental source: B cell
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-90/Domain: immunoglobulin homology <IMM>

Query Match
Best Local Similarity 92.1%; Score 82; DB 2; Length 110;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKODSEKRYVDSVKG 17
   |||||
Db 42 NIKODSEKRYVDSVKG 58
```

```
RESULT 5
S36280
Ig heavy chain V region (clone alpha-FOG1-A3) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
C:Accession: S36280
R:Griffiths, A.D.; Malnqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty,
EMBO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries
A:Reference number: S36256; MUID:93178448
A:Accession: S36280
A>Status: Preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-114 <GR1>
A:Cross-references: EMBL:218822
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
```

```
Query Match
Best Local Similarity 92.1%; Score 82; DB 2; Length 114;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKODSEKRYVDSVKG 17
   |||||
Db 50 NIKODSEKRYVDSVKG 66
```

```
RESULT 6
S78486
Ig heavy chain V region (clone FL13-28) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 21-Jan-2000
C:Accession: S78486; S31115
R:Raaphorst, F.M.
submitted to the EMBL Data Library, October 1991
A:Reference number: S78486
A:Accession: S78486
A:Molecule type: mRNA
A:Residues: 1-117 <RAA>
A:Cross-references: EMBL:X62965
A:Experimental source: clone FL13-28
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuur
Eur. J. Immunol. 22, 247-251, 1992
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comp
A:Reference number: S31104; MUID:92111633
A:Accession: S31115
A:Molecule type: mRNA
A:Residues: 1-4, 'L', 6-32, 'G', 34-52, 'E', 54-73, 'K', 75-97 <RAW>
A:Cross-references: EMBL:X62965
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
```

```
Query Match
Best Local Similarity 92.1%; Score 82; DB 2; Length 117;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NIKODSEKRYVDSVKG 17
   |||||
Db 50 NIKODSEKRYVDSVKG 66.
```

```
RESULT 7
S17079
Ig heavy chain V-gene (clone HKG19) - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C:Accession: S60299; S17079
```

R:Kuepers, R.; Fischer, U.; Rajewsky, K.; Gause, A.
Immunol. Lett. 34, 57-62, 1992
A:Title: Immunoglobulin heavy and light chain gene sequences of a human CD5 positive lmt
A:Reference number: S60295; MUID:93122853
A:Accession: S60299
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <RU2>
A:Cross-references: EMBL:X62128; NID:q38340; PIDN:CAA44059.1; PID:q38341
A>Note: the authors did not translate the codons for residues 6, 52, 54, 68, 69, 71, 72.
C:Genetics:
A:Introns: 16/1
C:Superfamily: immunoglobulin V region: immunoglobulin homology
C:Keywords: heterotetramer: immunoglobulin
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 92.1%; Score 82; DB 2; Length 117;
Best Local Similarity 88.2%; Pred. No. 1.4e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODSEKRYADSVRG 17
|||||
Db 69 NIKODSEKRYADSVRG 85

RESULT 8
S31675
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S31675
R:Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.
submitted to the EMBL data library, June 1992
A:Description: Mechanisms that generate human immunoglobulin diversity operate from the
A:Reference number: S31385
A:Accession: S31675
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-122 <CU2>
A:Cross-references: EMBL:Z14175; NID:q31015; PIDN:CAA78544.1; PID:q31016
C:Superfamily: Immunoglobulin V region: immunoglobulin homology
C:Keywords: heterotetramer: immunoglobulin
F:21-104/Domain: immunoglobulin homology <IMM>

Query Match 92.1%; Score 82; DB 2; Length 122;
Best Local Similarity 88.2%; Pred. No. 1.5e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODSEKRYADSVRG 17
|||||
Db 56 NIKODSEKRYADSVRG 72

RESULT 9
S31587
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S31587
R:Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.
submitted to the EMBL Data Library, June 1992
A:Description: Mechanisms that generate human immunoglobulin diversity operate from the
A:Reference number: S31585
A:Accession: S31587
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-136 <CU1>
A:Cross-references: EMBL:Z14189; NID:q31005; PIDN:CAA78558.1; PID:q31006
C:Superfamily: immunoglobulin V region: immunoglobulin homology
C:Keywords: heterotetramer: immunoglobulin
F:31-114/Domain: immunoglobulin homology <IMM>

Query Match 92.1%; Score 82; DB 2; Length 136;
Best Local Similarity 88.2%; Pred. No. 1.6e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODSEKRYADSVRG 17
|||||
Db 66 NIKODSEKRYADSVRG 82

RESULT 10
I37781
Ig variable region (VD) (clone T21-9) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
C:Accession: I37781; S25475
R:Demaison, C.; Chastagner, P.; Theze, J.; Zouali, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
A:Title: Somatic diversification in the heavy chain variable region genes expressed b
A:Reference number: A36876; MUID:94119917
A:Accession: I37781
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-139 <RES>
A:Cross-references: EMBL:X67908; NID:q33580; PIDN:CAA48106.1; PID:q33581
C:Superfamily: immunoglobulin V region: immunoglobulin homology
F:35-118/Domain: immunoglobulin homology <IMM>

Query Match 92.1%; Score 82; DB 2; Length 139;
Best Local Similarity 88.2%; Pred. No. 1.7e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODSEKRYADSVRG 17
|||||
Db 70 NIKODSEKRYADSVRG 86

RESULT 11
S22657
Ig heavy chain precursor V region (0-81VH) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 06-Feb-1998
C:Accession: S22657
R:Hirabayashi, Y.; Munakata, Y.; Sasaki, T.; Sano, H.
Nucleic Acids Res. 20, 2601, 1992
A:Title: Variable regions of a human anti-DNA antibody O-81 possessing lupus nephriti
A:Reference number: S22657; MUID:92285150
A:Accession: S22657
A:Molecule type: mRNA
A:Residues: 1-140 <HIR>
A:Cross-references: EMBL:X59134
C:Superfamily: Immunoglobulin V region: immunoglobulin homology
C:Keywords: heterotetramer: immunoglobulin
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-140/Product: Ig heavy chain (fragment) #status predicted <MAT>
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 88.8%; Score 79; DB 2; Length 140;
Best Local Similarity 82.4%; Pred. No. 5.4e-06;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODSEKRYADSVRG 17
|||||
Db 69 NVKODSARYADSVRG 85

RESULT 12
S31509
Ig heavy chain - human
C:Species: Homo sapiens (man)

Query Match 75.38; Score 67; DB 2; Length 136;

Search completed: November 19, 2001, 08:14:20
Job time: 193 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:33 ; Search time 27.32 Seconds
(Without alignments)
21.316 Million cell updates/sec

Title: US-09-610-118-62

Perfect score: 89
Sequence: 1 NIKQDSEKRYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	63	70.8	116 1 HV3T_HUMAN	P01781 homo sapien
2	57	64.0	122 1 HV3G_HUMAN	P01768 homo sapien
3	56	62.9	118 1 HV3V_HUMAN	P80419 homo sapien
4	55	61.8	119 1 HV3L_HUMAN	P01773 homo sapien
5	55	61.8	126 1 HV3K_HUMAN	P01772 homo sapien
6	50	56.2	119 1 HV3M_HUMAN	P01774 homo sapien
7	50	56.2	119 1 HV3N_HUMAN	P01775 homo sapien
8	49	55.1	119 1 HV3I_HUMAN	P01770 homo sapien
9	47	52.8	122 1 HV3A_HUMAN	P01762 homo sapien
10	47	52.8	122 1 HV3H_HUMAN	P01769 homo sapien
11	46	51.7	114 1 HV3B_HUMAN	P01763 homo sapien
12	46	51.7	117 1 HV03_CARAU	P19180 carassius a
13	46	51.7	121 1 HV3J_HUMAN	P01771 homo sapien
14	45	50.6	98 1 HV37_MOUSE	P18528 mus musculu
15	45	50.6	116 1 HV05_CARAU	P18181 carassius a
16	44	49.4	115 1 HV3F_HUMAN	P01767 homo sapien
17	44	49.4	421 1 YE04_METJA	Q58799 methanococ
18	43	48.3	117 1 HV3C_HUMAN	P01764 homo sapien
19	43	48.3	120 1 HV3E_HUMAN	P01766 homo sapien
20	42	47.2	117 1 HV53_MOUSE	P18524 mus musculu
21	42	47.2	258 1 YC09_YEAST	P25637 saccharomyc
22	41	46.1	117 1 HV01_CAICR	P01813 calman cicc
23	41	46.1	883 1 DYN_DROME	P27619 dirosophila
24	41	46.1	117 1 PKSK_BACSU	P40803 bacillus su
25	40.5	45.5	197 1 B1NL_STAAU	P18358 staphylococ
26	40	44.9	97 1 HV56_MOUSE	P18527 mus musculu
27	40	44.9	117 1 HV54_MOUSE	P18525 mus musculu
28	40	44.9	117 1 HV55_MOUSE	P18526 mus musculu
29	40	44.9	231 1 PSBQ_ONOVI	O22591 onobrychis
30	40	44.9	318 1 FMT_HAEIN	P44787 haemophilus
31	40	44.9	463 1 YCE8_YEAST	P25380 saccharomyc
32	39	43.8	113 1 HV3I_MOUSE	P01800 mus musculu
33	39	43.8	117 1 HV58_MOUSE	P18529 mus musculu

34	39	43.8	178 1 Y319_MYCGE
35	39	43.8	213 1 GPH_AQOAE
36	39	43.8	259 1 SRLD_ECOLI
37	39	43.8	426 1 GUNZ_ERWCH
38	39	43.8	432 1 ACDB_RAT
39	39	43.8	564 1 STP_BACSU
40	39	43.8	561 1 NERR_MOUSE
41	39	43.8	862 1 TP3B_MOUSE
42	38.5	43.3	117 1 HV02_CANFA
43	38	42.7	120 1 HV3U_HUMAN
44	38	42.7	149 1 NDK1_SACOF
45	38	42.7	159 1 GLBH_CAEEL

ALIGNMENTS

```

RESULT 1
HV3T_HUMAN STANDARD; PRT; 116 AA.
AC P01781;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION GAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP MEDLINE=75059123; PubMed=4803843;
RX Watanabe S., Barnikol H.U., Horn J., Berttram J., Hilschmann N.;
RT "The primary structure of a monoclonal IgM-immunoglobulin
RT (macroglobulin gal.), II: the amino acid sequence of the H-chain (mu-
RT type), subgroup H III. Architecture of the complete IgM-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
RN [2]
RP REVISION TO THE COMPOSITION OF 28-33.
RA Hilschmann N.;
RL Submitted (JUN-1975) to the PIR data bank.
CC -I- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC MACROGLOBULIN.
DR PIR: A02064; M3HUGL.
DR HSSP: P01772; 2IG2.
DR InterPro: IPR003006;
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9A9AAA1282 CRC64;

Query Match 70.8%; Score 63; DB 1; Length 116;
Best local similarity 58.8%; Pred. No. 0.0005;
Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 NIKQDSEKRYADSVRG 17
DB 50 NIKZBGSZBYVDYVKG 66

RESULT 2
HV3G_HUMAN STANDARD; PRT; 122 AA.
AC P01768;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION CAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;

```

RN [1]
 RP MEDLINE=81013859; PubMed=674332;
 RA Lehman D.W., Putnam F.W.;
 RT "Amino acid sequence of the variable region of a human mu chain;
 RL location of a possible JH segment."
 CC Proc. Natl. Acad. Sci. U.S.A. 77:3239-3243(1980).
 DR -1 MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM THE PLASMA OF A
 CC PATIENT WITH MACROGLOBULINEMIA.
 DR PIR: A02051; M3HUM.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g.1.
 KW Immunoglobulin V region.
 FT MOD_RES 1
 FT NON_TER 122 122
 SQ SEQUENCE 122 AA; 13668 MW; A42D0F17D252F1C2 CRC64;

Query Match 64.0%; Score 57; DB 1; Length 122;
 Best Local Similarity 56.2%; Pred. No. 0.0058;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKDGSSEKYVADSVRG 17
 DB 51 ISYGBBKRYABSVRG 66

RESULT 3
 HV3V_HUMAN STANDARD; PRT; 118 AA.
 AC P80419;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION GAR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=9525298; PubMed=7737190;
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
 RT "Characterization of the two unique human anti-flavin monooxal
 RL Eur. J. Biochem. 228:886-893(1995).
 DR HSSP: P01810; 2FBJ.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g.1.
 KW Immunoglobulin V region.
 FT NON_TER 118 118
 SQ SEQUENCE 118 AA; 13087 MW; 6C21D810ED1B6D1F CRC64;

Query Match. 62.9%; Score 56; DB 1; Length 118;
 Best Local Similarity 62.5%; Pred. No. 0.0083;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 NIKDGSSEKYVADSVR 16
 DB 50 NIRDDETEKFSVSVR 65

RESULT 4
 HV3L_HUMAN STANDARD; PRT; 119 AA.
 AC P01773;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION BUR.
 OS Homo sapiens (Human).
 FT MOD_RES 1
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 12981 MW; 12A709A75344D024 CRC64;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=79151016; PubMed=107164;
 RA Putnam F.W., Liu Y.-S.V., Low T.L.K.;
 RT "Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal
 RT IgA1 protease, digestion, Fab and Fc fragments, and the complete
 RT amino acid sequence of the alpha 1 heavy chain."
 RL J. Biol. Chem. 254:2865-2874(1979).
 DR PIR: A02056; A1HUBR.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g.1.
 KW Immunoglobulin V region; Glycoprotein.
 FT MOD_RES 1
 FT DISULFID 22 96
 FT CARBOHYD 28 28
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 12981 MW; 12A709A75344D024 CRC64;

Query Match 61.8%; Score 55; DB 1; Length 119;
 Best Local Similarity 68.8%; Pred. No. 0.013;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 IKDGSSEKYVADSVRG 17
 DB 51 ISYGSSEKYVADSVRG 66

RESULT 5
 HV3K_HUMAN STANDARD; PRT; 126 AA.
 AC P01772;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION KOL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=83289131; PubMed=688494;
 RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
 RT "Three-dimensional structure determination of antibodies. Primary
 RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=81072295; PubMed=7441755;
 RA Margart M., Deisenhofer J., Huber R., Palm W.;
 RT "Crystallographic refinement and atomic models of the intact
 RT immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A
 RL and 1.0-A resolution."
 RL J. Mol. Biol. 141:369-391(1980).
 DR PIR: A02055; G1HOKL.
 DR PDB: 2FB4; 12-JUL-89.
 DR PDB: 2IG2; 12-JUL-89.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g.1.
 KW Immunoglobulin V region; 3D-structure.
 FT MOD_RES 1
 FT DISULFID 22 96
 FT STRAND 105 110
 FT STRAND 3 7
 FT STRAND 11 12
 FT TURN 14 15
 FT STRAND 18 25
 FT HELIX 29 31

RESULT	ID	STANDARD;	PRT;	117 AA.
HV03_CARAU	HV03_CARAU			
AC	P19180.			
DT	01-NOV-1990	(Rel. 16, Created)		
DT	01-NOV-1990	(Rel. 16, Last sequence update)		
DT	15-JUL-1999	(Rel. 38, Last annotation update)		

```

DE IG HEAVY CHAIN V REGION 3 PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88144476; PubMed=3125551;
RA Wilson M.R., Middleton D., Warr G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure
RT and family relationships of two genes and a pseudogene in a teleost
RT fish."
RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J03616; AAA50807.1; -.
DR PIR: A28966; A28966.
DR HSSP: P01772; 2IG2.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
DR KX Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 3.
FT DOMAIN 20 49 FRAMEWORK 1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 55 68 FRAMEWORK 2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 86 117 FRAMEWORK 3.
FT DISULFD 41 115 BY SIMILARITY.
FT NON_TER 117 117
SO SEQUENCE 117 AA; 13220 MW; 512B625003FA6ECB CRC64;

Query Match 51.7%; Score 46; DB 1; Length 117;
Best Local Similarity 56.2%; Pred. No. 0.44;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 IKDGESEKYYADSVRG 17
| ||| : ||| : |
Db 70 IYDGSVDVSYADTVK 85

RESULT 13
ID HV3J_HUMAN STANDARD: PRT: 121 AA.
AC P01771;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION HIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=79124695; PubMed=420800;
RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;
RT "Amino acid sequence of the VH region of human myeloma
RT cryoimmunoglobulin IgG H1L."
RL Biochemistry 18:553-560(1979).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR: A02054; G1HGH1.
DR HSSP: P01772; 2IG2.

```

```

DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT NON_TER 121 121
SO SEQUENCE 121 AA; 13566 MW; 480FC53610EF5DAB CRC64;

Query Match 51.7%; Score 46; DB 1; Length 121;
Best Local Similarity 61.5%; Pred. No. 0.46;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 DGESEKYYADSVRG 17
: || || ||| : |
Db 54 NGSRTYGDVSVKG 66

RESULT 14
ID HV57_MOUSE STANDARD: PRT: 98 AA.
AC P18528;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 6.96.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/CJ;
RX MEDLINE=89279149; PubMed=2499654;
RA Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
RT "Early onset of somatic mutation in immunoglobulin VH genes during
RT the primary immune response."
RL J. Exp. Med. 169:2007-2019(1989).
CC -I- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
DR PIR: J70501; HVMS96.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
DR KX Immunoglobulin V region.
FT NON_TER 98 98
SO SEQUENCE 98 AA; 11007 MW; B8644F7F92BF95B CRC64;

Query Match 50.6%; Score 45; DB 1; Length 98;
Best Local Similarity 56.2%; Pred. No. 0.54;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 IKDGESEKYYADSVRG 17
| || || ||| : |
Db 51 ISDGSYYTPDSVKG 66

RESULT 15
ID HV05_CARAU STANDARD: PRT: 116 AA.
AC P19181;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 5A PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88144476; PubMed=3125551;
RA Wilson M.R., Middleton D., Warr G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure

```

RT and family relationships of two genes and a pseudogene in a teleost
 fish.
 RT fish.
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
 DR PIR: B28966; B28966.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region; signal.
 FT SIGNAL 1 19
 FT CHAIN 20 116 IG HEAVY CHAIN V REGION 5A.
 FT DOMAIN 20 49 FRAMEWORK 1.
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING 1.
 FT DOMAIN 55 68 FRAMEWORK 2.
 FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING 2.
 FT DOMAIN 85 116 FRAMEWORK 3.
 FT DISULFID 41 114 BY SIMILARITY.
 FT NON_TER 116 116
 SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;

Query Match 50.6%; Score 45; DB 1; length 116;
 Best Local Similarity 66.7%; Pred. No. 0.65;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 GSEKYYADSVRG 17
 DB 73 GGSTYYADSVK 84

Search completed: November 19, 2001, 08:25:34
 Job time: 722 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:00 ; Search time 81.19 Seconds
(without alignments)
27.703 Million cell updates/sec

Title: US-09-610-118-62
Perfect score: 89
Sequence: 1 NIKODSEKRYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phase:*
11: sp_plant:*
12: sp_protocent:*
13: sp_unclassified:*
14: sp_vertebrate:*
15: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	92.1	95	4 Q9ULB6	Q9ULB6 homo sapien
2	71	79.8	78	4 Q75730	Q75730 homo sapien
3	68	76.4	74	4 Q75744	Q75744 homo sapien
4	68	76.4	78	4 Q75720	Q75720 homo sapien
5	66	74.2	113	4 Q9UL90	Q9UL90 homo sapien
6	64	71.9	77	4 Q75726	Q75726 homo sapien
7	64	71.9	77	4 Q9UL93	Q9UL93 homo sapien
8	61	68.5	122	4 Q9UL84	Q9UL84 homo sapien
9	53	59.6	121	4 Q9UL71	Q9UL71 homo sapien
10	52	58.4	147	4 Q9Y509	Q9Y509 homo sapien
11	50	56.2	71	7 Q9GJ71	Q9GJ71 salmo trutt
12	49	55.1	80	4 Q75735	Q75735 homo sapien
13	48	53.9	118	4 Q9UL72	Q9UL72 homo sapien
14	48	53.9	594	2 Q9ZCY6	Q9ZCY6 rickettsia
15	47.5	53.4	235	5 Q9YZF9	Q9YZF9 drosophila
16	47	52.8	342	2 Q9EFK2	Q9EFK2 rhizobium t
17	46	51.7	984	2 Q67280	Q67280 aquifex aeo
18	45	50.6	77	4 Q75741	Q75741 homo sapien
19	45	50.6	150	4 Q9NQF2	Q9NQF2 homo sapien

20	45	50.6	169	4 Q9NQF1	Q9NQF1 homo sapien
21	45	50.6	430	2 Q31783	Q31783 bacillus su
22	44	49.4	71	7 Q9GJ72	Q9GJ72 salmo trutt
23	44	49.4	71	7 Q9GJ70	Q9GJ70 salmo trutt
24	44	49.4	71	7 Q9GJ69	Q9GJ69 salmo trutt
25	44	49.4	71	7 Q9GJ68	Q9GJ68 salmo trutt
26	44	49.4	71	7 Q9GJ67	Q9GJ67 salmo trutt
27	44	49.4	71	7 Q9GJ64	Q9GJ64 salmo trutt
28	44	49.4	140	4 Q9UNM8	Q9UNM8 homo sapien
29	44	49.4	472	4 Q9Y646	Q9Y646 homo sapien
30	44	49.4	541	4 Q9Y5X6	Q9Y5X6 homo sapien
31	43	48.3	71	7 Q31367	Q31367 brachydanio
32	43	48.3	71	7 Q9XR6	Q9XR6 salmo trutt
33	43	48.3	71	7 Q9GJ66	Q9GJ66 salmo trutt
34	43	48.3	71	7 Q9GJ65	Q9GJ65 salmo trutt
35	43	48.3	71	7 Q9GJ63	Q9GJ63 salmo trutt
36	43	48.3	71	7 Q9GJ62	Q9GJ62 salmo trutt
37	43	48.3	213	5 Q9VDN8	Q9VDN8 drosophila
38	43	48.3	317	2 Q30830	Q30830 rhodobacter
39	42	47.2	71	7 Q9XR9	Q9XR9 sigmodon hi
40	42	47.2	71	7 Q9XR5	Q9XR5 sigmodon hi
41	42	47.2	71	7 Q9XR3	Q9XR3 sigmodon hi
42	42	47.2	71	7 Q9GJ63	Q9GJ63 sigmodon hi
43	42	47.2	238	7 Q31467	Q31467 morone saxa
44	41.5	46.6	133	2 Q91656	Q91656 salmonella
45	41	46.1	82	4 Q75729	Q75729 homo sapien

ALIGNMENTS

RESULT 1
ID Q9ULB6 PRELIMINARY: PRT: 95 AA.
AC Q9ULB6:
DT 01-MAY-2000 (TREMUREL.13, Created)
DT 01-MAY-2000 (TREMUREL.13, Last sequence update)
DT 01-MAR-2001 (TREMUREL.16, Last annotation update)
DE IMONOGLOBULIN HEAVY CHAIN (FRAGMENT).
GN VH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tange Y., Kayano H.;
RT *Human VH gene sequence.*;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO IMONOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.
CC EMBL: AB035268; BAA87067.1; -.
DR HSSP: P01772; 2FE4.
DR InterPro: IPR003596; -.
DR InterPro: IPR003596; -.
DR Pfam: PF00047; 1g; 1.
DR SMART: SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 95
SQ SEQUENCE 95 AA; 10527 MW; 90A8C6D16D22574A CRC64;

Query Match 92.1%; Score 82; DB 4; Length 95;
Best local similarity 88.2%; Pred. No. 5.3e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODSEKRYADSVRG 17
|||||
Db 49 NIKODSEKRYADSVRG 65

RESULT 2
Q75730

ID 075730 PRELIMINARY: PRT: 78 AA.
AC 075730;
DT 01-NOV-1998 (TREMBLREL. 08, Created)
DT 01-NOV-1998 (TREMBLREL. 08, Last sequence update)
DE 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
GN VH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RA Fischer M., Kuipers R.;
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
RT mutated VH region genes.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ009527; CA08734.1; -
DR InterPro: IPR003596; -
DR SMART: SM00406; IGV: 1.
FT NON_TER 1 1
FT NON_TER 78 78
SQ SEQUENCE 78 AA: 8994 MW: D8C3975115720269 CRC64;

Query Match 79.8%; Score 71; DB 4; Length 78;
Best Local Similarity 76.5%; Pred. NO. 0.00025;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NIKDGSERYADSVRG 17
II :||| |||||:1
Db 21 NINEDSGKYYADSVKG 37
RESULT 3
075744 PRELIMINARY: PRT: 74 AA.
AC 075744;
DT 01-NOV-1998 (TREMBLREL. 08, Created)
DT 01-NOV-1998 (TREMBLREL. 08, Last sequence update)
DE 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
GN VH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RA Fischer M., Kuipers R.;
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
RT mutated VH region genes.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AJ009545; CA08748.1; -
DR InterPro: IPR003006; -
DR InterPro: IPR003596; -
DR Pfam: PF00047; IGV: 1.
DR SMART: SM00406; IGV: 1.
FT NON_TER 1 1
FT NON_TER 74 74
SQ SEQUENCE 74 AA: 8541 MW: EAF7DC54C84921DA CRC64;

Query Match 76.4%; Score 68; DB 4; Length 74;
Best Local Similarity 70.6%; Pred. NO. 0.00073;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY 1 NIKDGSERYADSVRG 17
II :||| |||||:1

Db 19 NINEDSGKYYADSVKG 35
RESULT 4
075720 PRELIMINARY: PRT: 78 AA.
AC 075720;
DT 01-NOV-1998 (TREMBLREL. 08, Created)
DT 01-NOV-1998 (TREMBLREL. 08, Last sequence update)
DE 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
GN VH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RA Fischer M., Kuipers R.;
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
RT mutated VH region genes.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ009517; CA08724.1; -
DR InterPro: IPR003596; -
DR SMART: SM00406; IGV: 1.
FT NON_TER 1 1
FT NON_TER 78 78
SQ SEQUENCE 78 AA: 9036 MW: BDD1797224418B17 CRC64;

Query Match 76.4%; Score 68; DB 4; Length 78;
Best Local Similarity 70.6%; Pred. NO. 0.00078;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NIKDGSERYADSVRG 17
II :||| |||||:1
Db 15 NIKHDSERYADSVKG 31
RESULT 5
09UL90 PRELIMINARY: PRT: 113 AA.
AC 09UL90;
DT 01-MAY-2000 (TREMBLREL. 13, Created)
DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
DE 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-9827113; Pubmed-9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Bernay S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AF035024; AAD56260.1; -
DR InterPro: IPR003006; -
DR InterPro: IPR003596; -
DR Pfam: PF00047; IGV: 1.
DR SMART: SM00406; IGV: 1.
FT NON_TER 1 1
FT NON_TER 113 113
SQ SEQUENCE 113 AA: 12437 MW: ED57FDD19086D07F CRC64;

Query Match 74.2%; Score 66; DB 4; Length 113;


```
QY      1 NIKQDGSEKYYADSVRG 17
        ||:|:|:|:|:|:|
Db     19 NIEQNGRERYSVDSKDG 35
```

RESULT 13
09UL72 PRELIMINARY: PRT: 118 AA.
ID 09UL72
AC 09UL72
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AF035042; AAD56278.1; -
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003006; -
DR InterPro: IPR003596; -
DR Pfam: PF00047; 19; 1.
DR SMART: SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 118
FT NON_TER 118
SQ SEQUENCE 118 AA; 12872 MW; BAD1A594AB2D5CCA CRC64;

Query Match 53.9%; Score 48; DB 4; Length 118;
Best Local Similarity 47.1%; Pred. No. 2.2;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NIKDSEKRYADSVRG 17
DB 49 SVTSGSSRYADSVRG 65

RESULT 14
09ZCY6 PRELIMINARY: PRT: 594 AA.
ID 09ZCY6
AC 09ZCY6
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE PENICILLIN-BINDING PROTEIN (PBP1).
GN PPS65.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=MADRID E.
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Scheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Neslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria";
RL Nature 396:133-140(1998).
RL EMBL: AJ235272; CAAL5013.1; -
DR InterPro: IPR001460; -
DR Pfam: PF00905; Transpeptidase; 1.
SQ SEQUENCE 594 AA; 67195 MW; 669AB06BAD4A28AC CRC64;

Query Match 53.9%; Score 48; DB 2; Length 594;
Best Local Similarity 52.9%; Pred. No. 14;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 NIKDSEKRYADSVRG 17
DB 194 NIKSGIEKRYADSVRG 210

RESULT 15
09VZFG PRELIMINARY: PRT: 235 AA.
ID 09VZFG
AC 09VZFG
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CG1259 PROTEIN.
GN CG1259.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abrahams J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkov A., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferrara C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RL EMBL: AE003480; AAF47864.1; -
DR Flybase: FBgn0035513; CG1259.
DR InterPro: IPR000618; -
DR Pfam: PF00379; Insect_cuticle; 1.
DR PRINTS: PR00947; CUTICLE.
DR PROSITE: PS00213; CUTICLE; 1.
SQ SEQUENCE 235 AA; 23066 MW; A96DC2E7914B5A68 CRC64;

Query Match 53.4%; Score 47.5; DB 5; Length 235;
Best Local Similarity 52.9%; Pred. No. 14;

Best Local Similarity 52.6%; Pred. No. 5.7;
Matches 10; Conservative 3; Mismatches

3; Indels

3; Gaps 1;

QY 2 IKODSEK---YYADSVRG 17

1:111:1111:1
Db 167 IEPOGSRIVSYADSLNG 185

Search completed: November 19, 2001, 08:25:01
Job time: 739 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:12:34 ; Search time 77.92 Seconds
(without alignments)
13.226 Million cell updates/sec

Title: US-09-610-118-62

Perfect score: 89
Sequence: 1 NIKQDSEKRYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%
Listing first 45 summaries

Database :

1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	22	AA61294 Anti-TANCO 268 scF
2	82	92.1	98	15	AA65206 Heavy chain variab
3	82	92.1	98	16	AA672074 DP54 VH region. H
4	82	92.1	98	21	AA640138 Anti-HIL12 antibod
5	82	92.1	98	21	AA640140 Anti-HIL12 antibod
6	82	92.1	117	16	AA66301 Human immunoglobul
7	82	92.1	125	22	AA662776 Human HIV-1 monocl
8	82	92.1	271	20	AA690180 Human clone A6 fus
9	82	92.1	312	20	AA683323 Single chain Apo-2
10	75	84.3	17	20	AA65049 Tumour antigen ant
11	73	82.0	17	20	AA65055 Tumour antigen ant

12	72	80.9	17	21	AA69824 Anti-HIL12 antibod
13	71	79.8	17	18	AA624547 CDR #2 of r101-2 h
14	71	79.8	130	18	AA624537 Immunoglobulin r10
15	69	77.5	17	21	AA639833 Anti-HIL12 antibod
16	69	77.5	116	16	AA66340 Human immunoglobul
17	67	75.3	17	22	AA639822 Anti-HIL12 antibod
18	67	75.3	123	22	AA62778 Human HIV-1 monocl
19	66	74.2	17	21	AA639487 Anti-HIL-12 antibo
20	66	74.2	17	21	AA639503 Anti-HIL-12 antibo
21	66	74.2	17	21	AA639511 Anti-HIL12 antibod
22	66	74.2	17	21	AA639819 Anti-HIL12 antibod
23	66	74.2	17	21	AA639820 Anti-HIL12 antibod
24	66	74.2	17	21	AA639821 Anti-HIL12 antibod
25	66	74.2	17	21	AA639823 Anti-HIL12 antibod
26	66	74.2	17	21	AA639825 Anti-HIL12 antibod
27	66	74.2	17	21	AA639826 Anti-HIL12 antibod
28	66	74.2	17	21	AA639827 Anti-HIL12 antibod
29	66	74.2	17	21	AA639828 Anti-HIL12 antibod
30	66	74.2	17	21	AA639834 Anti-HIL12 antibod
31	66	74.2	17	21	AA639842 Anti-HIL12 antibod
32	66	74.2	17	21	AA639865 Anti-HIL12 antibod
33	66	74.2	17	21	AA639868 Anti-HIL12 antibod
34	66	74.2	17	21	AA639870 Anti-HIL12 antibod
35	66	74.2	17	21	AA639871 Anti-HIL12 antibod
36	66	74.2	17	21	AA639881 Anti-HIL12 antibod
37	66	74.2	98	21	AA640123 Anti-HIL12 antibod
38	66	74.2	98	21	AA640141 Anti-HIL-12 antibo
39	66	74.2	114	21	AA639499 Anti-HIL-12 antibo
40	66	74.2	115	21	AA639507 Anti-HIL-12 antibo
41	66	74.2	115	21	AA639515 Anti-HIL12 antibody
42	66	74.2	115	21	AA639517 Anti-HIL12 antibody
43	66	74.2	115	21	AA639519 Anti-HIL12 antibody
44	66	74.2	115	21	AA639521 Anti-HIL12 antibody
45	66	74.2	115	21	AA639523 Anti-HIL12 antibody

ALIGNMENTS

RESULT 1		
ID	AA61294	standard; Peptide: 17 AA.
XX	AA61294;	
AC		
XX		
DT	04-APR-2001	(first entry)
XX		
DE	Anti-TANCO 268 scFv CDR, SEQ ID NO: 62.	
XX		
KW	Human; antibody; scFv; CDR; complementarity determining region;	
KW	TANCO 268; cardiant; cerebroprotective; cytosolic; anticoagulant;	
KW	thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;	
KW	platelet membrane glycoprotein receptor; bleeding disorder;	
KW	blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke	
KW	ischaemia; cardiovascular disease; immunological disease; liver disorder	
KW	cancer.	
OS	Homo sapiens.	
XX		
PN	WO200100810-A1.	
XX		
PD	04-JAN-2001.	
XX		
PF	30-JUN-2000; 2000WO-US18152.	
XX		
PR	30-JUN-1999; 99US-0345468.	
XX	06-DEC-1999; 99US-0454824.	
PR	14-FEB-2000; 2000US-0503387.	
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX	Busfield ST, Villalaj J, Jandrot-Petrus W, Valchencker W, Gill DS;	
PI	Qian MD, Kingsbury G;	

```

XX DR MPI: 2001-080877/09.
XX
PT New gene encoding human platelet-expressed collagen receptor,
PT glycoprotein VI, and its modulators, useful for preventing, treating
PT and diagnosing hemorrhagic disorders, thrombotic diseases and
XX immunological disorders -
XX
PS Claim 31: Page 102; 227pp; English.
XX
CC The present sequence is given in a specification relating to an isolated
CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC glycoprotein VI (GPVI), also called RANGO 268. The GPVI polynucleotides
CC and polypeptides and their modulators, e.g. antisense nucleic acids,
CC ribozymes and antibodies, are useful for preventing, treating and
CC diagnosing disorders associated with aberrant expression or activity of
CC GPVI. These disorders include bleeding disorders
CC (e.g. thrombocytopaenia), blood vessel injury, thrombotic disorders
CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC infarction), immunological diseases (e.g. platelet disorder) and
CC embryonic liver disorders. Preferably they are used to prevent acture
CC cardiac ischaemia following angioplasty and metastatic cancers,
CC especially of the colon and liver.
XX
SO Sequence 17 AA:

```

Query Match 100.0%: Score 89: DB 22: Length 17;
 Best Local Similarity 100.0%: Pred. No. 1.1e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 NIKQDSEKRYADSVRG 17
Db 1 nlkqdgsekryadvsvrg 17

```

RESULT 2
 AAR52066 standard; Protein: 98 AA.

AAR52066:
 11-OCT-1996 (first entry)

Heavy chain variable region of human PL0123 antibody.

antibody; humanised; murine; human; heavy chain; light; variable;
 framework region; complementarity determining region; reshaping;
 modelling; surface residue; modity.

Homo sapiens.

Key location/Qualifiers
 Region 1..30 /label=framework_region_1
 Region 31..35 /note="FR 1"
 Region 36..49 /label=complementarity_determining_region_1
 Region 50..59 /note="FR 2"
 Region 60..98 /note="CDR 2"
 Region /note="FR 3"

EP592106-A1.
 13-APR-1994.
 07-SEP-1993; 93EP-0307051.

```

XX PR 09-SEP-1992; 92US-0942245.
XX
PA (PEDE/) PEDERSEN J T.
PA (IMMU-) IMMUNOGEN INC.
XX
PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
XX MPI: 1994-120230/15.
XX
PT Method of resurfacing of rodent antibodies to produce humanised
PT antibody forms - for producing non-human antibodies with improved
PT therapeutic efficiency by presenting human surface on V-region
XX
PS Example 1: Fig 4B; 230pp; English.
XX
CC Modification of a rodent antibody (Ab) or fragment by resurfacing in
CC order to produce a humanised rodent Ab can be determined by calculating
CC homology between murine and human Ab antibody surfaces. In order to test
CC the resurfacing approach of the invention, three humanisation
CC experiments were set up: (1) traditional loop grafting; (2) resurfacing
CC approach using most similar chain; and (3) resurfacing approach using
CC human sequences with most similar surface residues. The Ab used was the
CC murine anti-N901 Ab (see AAR52061). Experiment 3 was carried out using
CC the present sequence which represents the human PL0123 Ab light chain
CC variable region with 74 percent homology with anti-N901 Ab. N901/PL0123
CC (AAR52067) was prepd. by resurfacing. Sequence numbering starts at
XX position 118 in the specification.
XX
SO Sequence 98 AA:

```

Query Match 92.1%: Score 82: DB 15: Length 98;
 Best Local Similarity 88.2%: Pred. No. 1.2e-06;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 NIKQDSEKRYADSVRG 17
Db 50 nlkqdgsekryadvsvkg 66

```

RESULT 3
 AAR72074 standard; Protein: 98 AA.

AAR72074:
 26-SEP-1995 (first entry)

DP54 VH region.

Graves ophthalmopathy associated immunoglobulin protein;
 orbital antigen; monoclonal antibody; heavy chain; H chain;
 variable region; autoimmunity.

Homo sapiens.

Key location/Qualifiers
 Region 41..44 /label=CDR1
 Region 49..67 /label=CDR2
 WO9508336-A.
 30-MAR-1995.
 22-SEP-1994; 94WO-US10756.
 22-SEP-1993; 93US-0124469.
 (NICH-) NICHOLS INST DIAGNOSTICS.

PI mclachlan SM, Rapoport B;
XX
DR WPI: 1995-139383/18.
DR N-PSDB: AA089332.
XX
PT 'Graves' ophthalmopathy-associated monoclonal antibody - produced
PT by molecular cloning of immunoglobulin genes by PCR
XX
PS Disclosure; Page 72; 94pp; English.
XX
CC L- and H-chain DNA was amplified by PCR from Graves' orbital
CC tissue and clones encoding autoimmune-associated immunoglobulin
CC fragments were obtained. 2/15 clones of H chain (IgG1) genes
CC showed homology to the germ-line gene DP54 (AA089332), which encodes
CC the protein given in AAR72074. The DNA (AA089333) and corresp. amino
CC acid (AAR72075) sequences of the VH region of a representative clone,
CC OF7H1.19, are provided.
CC
XX
SQ Sequence 98 AA;

Query Match 92.1%; Score 82; DB 16; Length 98;
Best Local Similarity 88.2%; Pred. NO. 1.2e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODGSEKYYADSVRG 17
|||||
DB 50 nikdgsekyvdsvk 66

RESULT 4
AAB40138
ID AAB40138 standard; Protein: 98 AA.
XX
AC AAB40138;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-IL12 antibody H chain V region amino acid sequence SEQ ID 664.
XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarily determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiatic;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PF 24-MAR-2000; 2000WO-US07946.
XX
PR 25-MAR-1999; 99US-0126603.
XX
PA (BADI) BASF AG.
PA (GENY) GENETICS INST INC.
XX
XX Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
DR WPI: 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
XX
PS Claim 75; Page 122; 377pp; English.
XX
CC This invention relates to a new human antibody specific for human

CC interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarily determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
CC represent anti-IL-12 CDR3 related amino acid sequences. AAB39772-B40063
CC represent other CDR sequences. Light chain CDR3 consensus sequences are
CC given in AAB40064-B40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in
CC AAC61062-C61071. The antibody of the invention is a neutralising
CC antibody and has antirheumatic; antiarthritic; antisclerotic;
CC antiinflammatory; neuroprotective; antipsoriatic; antiasthmatic;
CC cardiatic; antiparasitic; antibacterial and immunosuppressive activity.
CC The antibodies or antigen-binding fragments are useful in the treatment
CC of disorders associated with detrimental release of human IL-12,
CC especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
CC They can also be used in the manufacture of a pharmaceutical composition
CC to treat human IL-12 disorders.
CC
XX
SQ Sequence 98 AA;

Query Match 92.1%; Score 82; DB 21; Length 98;
Best Local Similarity 88.2%; Pred. NO. 1.2e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODGSEKYYADSVRG 17
|||||
DB 50 nikdgsekyvdsvk 66

RESULT 5
AAB40140
ID AAB40140 standard; Protein: 98 AA.
XX
AC AAB40140;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-IL12 antibody H chain V region amino acid sequence SEQ ID 666.
XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarily determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiatic;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200056772-A1.
XX
PD 28-SEP-2000.
XX
PF 24-MAR-2000; 2000WO-US07946.
XX
PR 25-MAR-1999; 99US-0126603.
XX
PA (BADI) BASF AG.
PA (GENY) GENETICS INST INC.
XX
XX Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
DR WPI: 2000-638250/61.
XX
PT New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis -
XX
PS Claim 75; Page 122; 377pp; English.
XX

XX This invention relates to a new human antibody specific for human
CC Interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarity determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
CC represent anti-IL-12 CDR3 related amino acid sequences. AAB39772-B40063
CC represent other CDR sequences. Light chain CDR3 consensus sequences are
CC given in AAB40064-B40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in
CC AAC61062-C61071. The antibody of the invention is a neutralising
CC antibody and has antirheumatic; antiarthritic; antisclerotic;
CC antiinflammatory; neuroprotective; antipsoriatic; antitumoric;
CC cardiant; antiparasitic; antibacterial and immunosuppressive activity.
CC The antibodies or antigen-binding fragments are useful in the treatment
CC of disorders associated with detrimental release of human IL-12,
CC especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
CC They can also be used in the manufacture of a pharmaceutical composition
CC to treat human IL-12 disorders.

SO Sequence 98 AA;

Query Match 92.1%; Score 82; DB 21; Length 98;
Best Local Similarity 88.2%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODGSEKRYADSVRG 17
|||
Db 50 nlkqdgsekryadvsvkg 66

RESULT 6
AAB66301
ID AAB66301 standard; Protein; 117 AA.

XX AAB66301;

DT 02-AUG-1995 (first entry)

DE Human immunoglobulin variable heavy chain #7.

KW primer; PCR; amplify; human; immunoglobulin; variable; heavy chain;

KM cosmid; placenta; vector; pJ8B1; E.coli; mammalian.

OS Homo sapiens.

XX WO9426895-A.

PD 24-NOV-1994.

PF 10-MAY-1993; 93WO-JP00603.

PR 10-MAY-1993; 93WO-JP00603.

PA (NTSB) JAPAN TOBACCO INC.

PI Honjo T, Matsuda F;

DR WPI: 1995-006791/01.

DR N-PSDB; AAO78946.

PT DNA fragment comprising human immunoglobulin Vh genes - for the
production of human immunoglobulin in mammalian hosts

PS Claim 17; Page 40-41; 130pp; Japanese.

XX Protein sequences (AAB66295-51) are novel human immunoglobulin heavy
chain sequences encoded by novel isolated genes. The genes
CC (AAO78939-79002) were isolated and cloned from a series of cosmid
CC constructs: Y202; Y103; Y21; Y6; Y24; 3-31; M84; M18 and M131, by PCR

CC amplification using primers AAO78917-38. The genes are subdivided into 5
CC families of Vh genes. The fragments cover a region of 800 kb. The DNA
CC fragments were isolated from high molecular weight DNA from human
CC placenta. The DNA was partially digested with *Bgl*I restriction enzyme.
CC The fragments were separated by gel electrophoresis and 35-45 kb fractions
CC were collected. The fragments were ligated with *Clai*I-digested cosmid
CC vector pJ8B1. The ligation products were in vitro packed and infected
CC into E.coli 490A. The fragments were then subcloned by colony
CC hybridisation. The Vh genes and the DNA fragments encoding them are
CC useful in producing human immunoglobulin in mammalian hosts.

SO Sequence 117 AA;

Query Match 92.1%; Score 82; DB 16; Length 117;
Best Local Similarity 88.2%; Pred. No. 1.5e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODGSEKRYADSVRG 17
|||
Db 69 nlkqdgsekryadvsvkg 85

RESULT 7

AAB62776
ID AAB62776 standard; Protein; 125 AA.

XX AAB62776;

DT 03-APR-2001 (first entry)

DE Human HIV-1 monoclonal antibody SEQ ID NO: 75.

KW Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;

KM envelope glycoprotein; gp120; diagnosis.

OS Homo sapiens.

XX WO200100678-A1.

PD 04-JAN-2001.

PF 23-JUN-2000; 2000WO-US17327.

PR 30-JUN-1999; 99US-0141701.

PA (USSR) US DEPT HEALTH & HUMAN SERVICES.

PI Watkins BA, Reitz MS;

DR WPI: 2001-112438/12.

DR N-PSDB; AAF29077.

PT Novel human monoclonal antibody immunoreactive with human
immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
in biological sample and providing passive immunotherapy to HIV-1
infected mammal

PS Claim 1; Page 69-70; 81pp; English.

XX The present invention provides the protein and coding sequences for the
CC variable regions of human monoclonal antibodies which are immunoreactive
CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
CC These can be used in diagnosis and therapy of HIV-1 infection.

SO Sequence 125 AA;

Query Match 92.1%; Score 82; DB 22; Length 125;
Best Local Similarity 88.2%; Pred. No. 1.7e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKODGSEKRYADSVRG 17

Db 51 nikqdgsekyyvsvkg 67

RESULT 8

AAW90180 standard; Protein; 271 AA.

AAW90180;

10-MAY-1999 (first entry)

Human clone A6 fusion protein antibody.

Antibody; human; clone A6; hab: fusion protein; polyclonal tag;

Western blot; enzyme-linked immunosorbent assay; ELISA; therapy;

immunofluorescence; immunoprecipitation assay; affinity purification;

diagnosis; vaccine; serum; immune response.

Homo sapiens.

Key Location/Qualifiers

Protein 1..271

/note="Partial coding sequence, no stop codon given"

WO901475-A2.

14-JAN-1999.

03-JUL-1998; 98WO-DE01882.

04-JUL-1997; 97DE-1028697.

(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

Braunagel M, Doersam H, Kipriyanov S, Kuerschner T;

Little M, Welschof M;

WPI: 1999-106000/09.

N-PSDB: AAV74274.

Human antibody against fusion protein with polyclonal tag -

useful as standard in immunoassays, for affinity purification,

diagnosis and therapy and for preparing vaccines

Claim 2; Fig 4; 20pp; German.

This invention describes a human antibody (hab) against a fusion

(poly)peptide or protein that includes a segment of at least 6

consecutive His residues. This antibody is useful in Western blots,

enzyme-linked immunosorbent assay (ELISA), immunofluorescence or

immunoprecipitation assays. Also hab can be used for affinity

purification of the protein, for in vivo diagnosis or therapy, and

in production of vaccines. hab are universally applicable

alternatives to human serum. They are specific for the polyclonal

tag, regardless of the nature of the rest of the protein. Since hab

are not produced in an animal, they contain no components that can

induce an immune response in humans.

Sequence 271 AA:

Query Match 92.1%; Score 82; DB 20; Length 271;

Best Local Similarity 88.2%; Pred. No. 4e-06;

Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKQDGSEKYYVDSVKG 17

Db 52 nikqdgsekyyvsvkg 68

RESULT 9

AAW83323

ID AAW83323 standard; Protein; 312 AA.

AAW83323;

16-MAR-1999 (first entry)

Single chain Apo-2 antibody 20B6.

Human; Apo-2; receptor; apoptosis; neurodegenerative disease; cancer;

tumour necrosis factor; TNF; tumour necrosis factor receptor; TNFR;

TNF cytokine.

Homo sapiens.

WO9851793-A1.

19-NOV-1998.

14-MAY-1998; 98WO-US09704.

09-FEB-1998; 98US-0020746.

15-MAY-1997; 97US-0857216.

(GETH) GENENTECH INC.

Adams CW, Ashkenazi AJ, Chuntharapai A, Kim KJ;

WPI: 1999-045228/04.

N-PSDB: AAV72533.

Human Apo-2 polypeptide inducing apoptosis - useful to treat

conditions linked with decreased apoptosis e.g. cancer, and produce

antibodies to increase or decrease apoptosis

Example 14; Fig 16; 134pp; English.

The present invention describes human Apo-2. Apo-2 can be used

therapeutically to induce apoptosis in mammalian cells, and so is useful

to treat conditions associated with decreased apoptosis e.g. cancer.

Apo-2 is believed to be a new tumour necrosis factor (TNF) receptor

(TNFR). TNF cytokines can induce apoptosis, thought to be initiated by

binding to TNFRs, and Apo-2 triggered caspase-dependent apoptosis. It

can be used to identify agents activating Apo-2, useful to treat

mammalian cancer cells, and to produce Apo-2 chimeras useful

therapeutically (e.g. those containing immunoglobulin sequences can be

inhibit apoptosis) or diagnostically (e.g. those comprising an epitope

tag polypeptide allow Apo-2 detection and purification using anti-tag

antibodies). It can be used to produce antibodies which can be combined

with a (particularly pharmaceutically acceptable) carrier in compositions

or used to produce dimeric molecules (especially homodimeric molecules

comprising first and second Apo-2 antibodies). Agonistic (especially

single-chain) antibodies can be administered to induce apoptosis in

mammalian cancer cells, and antagonistic antibodies used to block

excessive apoptosis (e.g. in neurodegenerative diseases). Apo-2

antibodies may also be used diagnostically e.g. to detect Apo-2

expression in cells/tissues and in Apo-2 purification. The present

sequence represents a single chain Apo-2 antibody, designated 20B6.

Sequence 312 AA:

Query Match 92.1%; Score 82; DB 20; Length 312;

Best Local Similarity 88.2%; Pred. No. 4.8e-06;

Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKQDGSEKYYVDSVKG 17

Db 89 nikqdgsekyyvsvkg 105

RESULT 10

AAV05049

ID AAV05049 standard; peptide; 17 AA.

```

XX AC AAY05049;
XX XX
XX DT 16-JUN-1999 (first entry)
XX XX
XX DE Tumour antigen antibody heavy chain CDR2 clone F15.
XX XX
XX KW Tumour antigen; antibody; CDR; complementarily determining region;
XX KW binding molecule identification; tumour-specific binding polypeptide;
XX KW cancer therapy; heavy chain.
XX OS Homo sapiens.
XX PN WO9906834-A2.
XX PD 11-FEB-1999.
XX PF 04-AUG-1998; 98WO-US16280.
XX PR 04-AUG-1997; 97US-0905825.
XX PR 04-AUG-1997; 97US-1112222.
XX PA (IXSY-) IXSYS INC.
XX PI Huse WD, Watkins JD, Wu H;
XX DR WPI: 1999-153951/13.
XX DR N-PSDB: AAX28226.
XX PS Identifying binding molecules for ligands, particularly tumour
XX PT antigens - by selectively immobilising a population of binding
XX PT molecules to a solid support and screening for binding to two or
XX PT more ligands
XX PS Claim 15; Page 60; 80pp; English.
XX XX
XX CC This sequence represents a heavy chain complementarily determining
XX CC region (CDR) from a tumour antigen specific antibody.
XX CC The invention relates to a method for identifying a binding molecule
XX CC having selective affinity for a ligand comprising: (a) selectively
XX CC immobilising a diverse population of binding molecules to a solid
XX CC support; (b) simultaneously contacting the diverse population immobilised
XX CC on the solid support with 2 or more ligands; and (c) determining at least
XX CC one binding molecule which selectively binds to one or more of the
XX CC ligands. The method allows for the rapid and efficient methods for the
XX CC identification of binding molecules which exhibit selective affinity for
XX CC one or more ligands of interest. They are used particularly for
XX CC identifying tumour-specific binding polypeptides which can be used as
XX CC targeting agents for cancer therapy that minimises impact on non-tumour
XX CC tissues.
XX SQ Sequence 17 AA:

```

```

Query Match 84.3%; Score 75; DB 20; Length 17;
Best Local Similarity 82.4%; Pred. No. 2.5e-06;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 1 NIKDGSSEKYYADSVRG 17
Db 1 nlnqdsdkyyvdsxg 17

```

```

RESULT 11
AAY05055
ID AAY05055 standard; peptide: 17 AA.
XX AC AAY05055;
XX DT 16-JUN-1999 (first entry)
XX DE Tumour antigen antibody heavy chain CDR2 clone F133.
XX XX

```

```

KW KW Tumour antigen; antibody; CDR; complementarily determining region;
KW KW binding molecule identification; tumour-specific binding polypeptide;
KW KW cancer therapy; heavy chain.
XX OS Homo sapiens.
XX PN WO9906834-A2.
XX PD 11-FEB-1999.
XX PF 04-AUG-1998; 98WO-US16280.
XX PR 04-AUG-1997; 97US-0905825.
XX PR 04-AUG-1997; 97US-1112222.
XX PA (IXSY-) IXSYS INC.
XX PI Huse WD, Watkins JD, Wu H;
XX DR WPI: 1999-153951/13.
XX DR N-PSDB: AAX28232.
XX PS Identifying binding molecules for ligands, particularly tumour
XX PT antigens - by selectively immobilising a population of binding
XX PT molecules to a solid support and screening for binding to two or
XX PT more ligands
XX PS Claim 15; Page 60; 80pp; English.
XX XX
XX CC This sequence represents a heavy chain complementarily determining
XX CC region (CDR) from a tumour antigen specific antibody.
XX CC The invention relates to a method for identifying a binding molecule
XX CC having selective affinity for a ligand comprising: (a) selectively
XX CC immobilising a diverse population of binding molecules to a solid
XX CC support; (b) simultaneously contacting the diverse population immobilised
XX CC on the solid support with 2 or more ligands; and (c) determining at least
XX CC one binding molecule which selectively binds to one or more of the
XX CC ligands. The method allows for the rapid and efficient methods for the
XX CC identification of binding molecules which exhibit selective affinity for
XX CC one or more ligands of interest. They are used particularly for
XX CC identifying tumour-specific binding polypeptides which can be used as
XX CC targeting agents for cancer therapy that minimises impact on non-tumour
XX CC tissues.
XX SQ Sequence 17 AA:

```

```

Query Match 82.0%; Score 73; DB 20; Length 17;
Best Local Similarity 76.5%; Pred. No. 5.6e-06;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 1 NIKDGSSEKYYADSVRG 17
Db 1 nlnqdsdkyyvdsxg 17

```

```

RESULT 12
AAB39824
ID AAB39824 standard; peptide: 17 AA.
XX AC AAB39824;
XX DT 05-FEB-2001 (first entry)
XX DE Anti-IL12 antibody heavy chain CDR2 amino acid sequence SEQ ID 340.
XX XX
XX KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
XX KW complementarily determining region; CDR; antirheumatic; antiarthritis;
XX KW antisclerotic; neuroprotective; antiparasitic; antisthmatic; cardiant;
XX KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
XX KW multiple sclerosis; rheumatoid arthritis.
XX OS Homo sapiens.

```

XX MO200056772-A1.
XX
XX 28-SEP-2000.
PD
XX 24-MAR-2000; 2000MO-US07946.
XX
XX 25-MAR-1999; 99US-0126603.
XX
XX (BADI) BASF AG.
PA (GEMV) GENETICS INST INC.
XX
XX Salfield JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M,
PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Wildon A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX
XX WPI: 2000-638250/61.
DR
XX
XX New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis.
XX
XX Claim 33; Figure 2B; 377pp; English.
XX
XX This invention relates to a new human antibody specific for human
CC interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarity determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
CC represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063
CC represent other CDR sequences, light chain CDR3 consensus sequences are
CC given in AAB40064-B40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in
CC AAB61062-C61071. The antibody of the invention is a neutralizing
CC antibody and has antirheumatic; antiarthritic; antisclerotic;
CC antiinflammatory; neuroprotective; antibacterial; immunosuppressive activity.
CC cardiant; antiparasitic; antitubercular and immunosuppressive activity.
CC The antibodies or antigen-binding fragments are useful in the treatment
CC of disorders associated with detrimental release of human IL-12,
CC especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
CC They can also be used in the manufacture of a pharmaceutical composition
CC to treat human IL-12 disorders.
XX
XX Sequence 17 AA;
SQ

Query Match 80.9%; Score 72; DB 21; Length 17;
Best Local Similarity 76.5%; Pred. No. 8.2e-06;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NIKDGSSEKYYADSVRG 17
II: ||| ||||| :|
Db 1 nlydgsnkyvdsyvk 17

RESULT 13
AAW24547
ID AAW24547 standard; peptide: 17 AA.
XX
XX AAW24547;
AC
XX
XX 09-OCT-1997 (first entry)
DT
XX
XX CDR #2 of r101-2 heavy chain variable region.
DE
XX
XX Immunoglobulin; variable region; heavy chain; thyrotropin receptor;
KW thyroid stimulating activity; light chain; Basedow's disease; antibody;
KW peripheral blood lymphocyte; complementarity determining region; CDR.
XX
XX Homo sapiens.
OS

XX JP09140386-A.
XX
XX 03-JUN-1997.
PD
XX 22-NOV-1995; 95JP-0328235.
XX
XX 22-NOV-1995; 95JP-0328235.
XX
XX (EIKE) EIKEN KAGAKU KK.
PA
XX
XX WPI: 1997-344899/32.
DR
XX
XX Antibody containing immunoglobulin heavy chain mutation - with
PT thyroid function stimulating activity
PT
XX
XX Claim 12; Page 14; 18pp; Japanese.
PS
XX
XX AAW24540-W24551 represent the complementarity determining regions (CDRs)
CC of immunoglobulin heavy and light chain variable regions isolated from
CC peripheral blood lymphocyte strains (see AAW24536-W24539 for full length
CC immunoglobulin chains). These sequences were isolated from the B6B7 and
CC 101-2 strains of peripheral blood lymphocytes of a Basedow's disease
CC patient. These sequences are replaced, deleted or inserted into an
CC antibody, to create the antibodies of the invention. The antibodies of
CC the invention have thyroid function stimulating activity, and act by
CC combining with thyrotropin receptor. The antibody can be used in a
CC method to detect autoantibodies which have thyroid function stimulating
CC activity.
XX
XX Sequence 17 AA;
SQ

Query Match 79.8%; Score 71; DB 18; Length 17;
Best Local Similarity 76.5%; Pred. No. 1.2e-05;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NIKDGSSEKYYADSVRG 17
II: ||| ||||| ||| :|
Db 1 dkpdsgekyvdsyvk 17

RESULT 14
AAW24537
ID AAW24537 standard; Protein; 130 AA.
XX
XX AAW24537;
AC
XX
XX 09-OCT-1997 (first entry)
DT
XX
XX Immunoglobulin r101-2 heavy chain variable region.
DE
XX
XX Immunoglobulin; variable region; heavy chain; thyrotropin receptor;
KW thyroid stimulating activity; light chain; Basedow's disease; antibody;
KW peripheral blood lymphocyte.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..19 /note="signal peptide"
FT Protein 20..130
XX
XX JP09140386-A.
XX
XX 03-JUN-1997.
PD
XX
XX 22-NOV-1995; 95JP-0328235.
XX
XX 22-NOV-1995; 95JP-0328235.
XX
XX (EIKE) EIKEN KAGAKU KK.
PA
XX

DR WPI: 1997-344899/32.
 DR N-PSDB; AAT79919.
 XX
 PT Antibody containing immunoglobulin heavy chain mutation - with
 PT thyroid function stimulating activity
 XX
 PS Claim 31: Page 15-16; 18pp; Japanese.
 XX
 CC AAM24536-W24539 represent the immunoglobulin heavy and light chain
 CC variable regions isolated from peripheral blood lymphocyte strains.
 CC These sequences were isolated from the B6B7 and 101-2 strains of
 CC peripheral blood lymphocytes of a Basedow's disease patient. These
 CC sequences are replaced, deleted or inserted into an antibody, to create
 CC the antibodies of the invention. The antibodies of the invention have
 CC thyroid function stimulating activity, and act by combining with
 CC thyrotropin receptor. The antibody can be used in a method to detect
 CC autoantibodies which have thyroid function stimulating activity.
 XX
 SQ Sequence 130 AA:

Query Match 79.8%; Score 71; DB 18; Length 130;
 Best Local Similarity 76.5%; Pred. No. 0.00013;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NIKDSEKYYADSVRG 17
 :||| ||||| |||||
 DB 69 dlkpdssekyydvsvkg 85

RESULT 15

AAB39833
 ID AAB39833 standard; Peptide; 17 AA.

AC AAB39833;

DT 05-FEB-2001 (first entry)

DE AnL1-hiL12 antibody heavy chain CDR2 amino acid sequence SEQ ID 349.

XX Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
 KW complementarity determining region; CDR; antirheumatic; antiarthritic;
 KW antisclerotic; neuroprotective; antipsoriatic; antiasclerotic; cardiant;
 KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis.

XX Homo sapiens.

OS WO200056772-A1.

PN 28-SEP-2000.

XX 24-MAR-2000; 2000WO-US07946.

XX 25-MAR-1999; 99US-0126603.

XX (BADI) BASF AG.

PA (GEMT) GENETICS INST INC.

XX Saifeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M,
 PI Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedlich S, Myles A;
 PI Vaidman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
 PI Derbyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
 DR WPI: 2000-638250/61.

XX New human antibody specific for human interleukin-12 (IL-12) used to
 PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
 PT disease and multiple sclerosis -

XX Claim 33; Figure 2B; 377pp; English.

XX This invention relates to a new human antibody specific for human

CC interleukin-12 (IL-12). The invention also includes antigen binding
 CC portions that bind to IL-12. Sequences AAB39485-B39516 represent human
 CC anti-IL-12 antibody heavy and light chain complementarity determining
 CC region (CDR) amino acid sequences, and also includes variable region
 CC amino acid sequences. Other variable region amino acid sequences are
 CC given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
 CC represent anti-IL-12 CDR3 related amino acid sequences. AAB39772-B40063
 CC represent other CDR sequences. Light chain CDR3 consensus sequences are
 CC given in AAB40064-B40067. Primers used in the identification and
 CC construction of the antibodies of the invention are given in
 CC AAC61062-C61071. The antibody of the invention is a neutralising
 CC antibody and has antirheumatic; antiarthritic; antisclerotic;
 CC antiinflammatory; neuroprotective; antipsoriatic; antiasclerotic;
 CC cardiant; antiparasitic; antibacterial and immunosuppressive activity.
 CC The antibodies or antigen-binding fragments are useful in the treatment
 CC of disorders associated with detrimental release of human IL-12,
 CC especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
 CC They can also be used in the manufacture of a pharmaceutical composition
 CC to treat human IL-12 disorders.

XX
 SQ Sequence 17 AA:

Query Match 77.5%; Score 69; DB 21; Length 17;
 Best Local Similarity 81.2%; Pred. No. 2.6e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKDSEKYYADSVRG 17
 :||| ||||| |||||
 DB 2 lkydgsnkyyadvsvkg 17

Search completed: November 19, 2001, 08:12:34
 Job time: 87 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:14 ; Search time 526.89 seconds
(without alignments)
8.959 Million cell updates/sec

Title: US-09-610-118-62

Perfect score: 89

Sequence: 1 NIKDGSSEKYYADSVRG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep: *
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep: *
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep: *
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep: *
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep: *
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep: *
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep: *
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep: *
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep: *
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep: *
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep: *
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep: *
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep: *
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep: *
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep: *
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep: *
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep: *
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep: *
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep: *
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep: *
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep: *
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep: *
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep: *
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	89	100.0	17	1 PCT-US00-18152-62
2	89	100.0	17	20 US-09-610-118-62
3	89	100.0	17	22 US-09-832-312-62
4	82	92.1	98	1 PCT-US00-07946-664
5	82	92.1	98	1 PCT-US00-07946-666
6	82	92.1	98	5 US-08-124-469-32
7	82	92.1	98	8 US-08-472-771-32
8	82	92.1	98	8 US-08-486-202-32
9	82	92.1	98	19 US-09-534-717-664

10	82	92.1	98	19	US-09-534-717-666	Sequence 666, App
11	82	92.1	98	21	US-09-751-181-85	Sequence 85, App1
12	82	92.1	124	11	US-08-712-564-10	Sequence 10, App1
13	82	92.1	125	24	US-60-141-701-75	Sequence 75, App1
14	82	92.1	134	22	US-09-811-737-3	Sequence 3, App1
15	82	92.1	241	1	PCT-US01-19110-1889	Sequence 1889, App
16	82	92.1	241	22	US-09-880-748-1889	Sequence 1889, App
17	82	92.1	247	1	PCT-US01-19110-1470	Sequence 1470, App
18	82	92.1	247	22	US-09-880-748-1470	Sequence 1470, App
19	82	92.1	251	1	PCT-US01-19110-1594	Sequence 1594, App
20	82	92.1	251	22	US-09-880-748-1594	Sequence 1594, App
21	82	92.1	253	1	PCT-US01-19110-1003	Sequence 1003, App
22	82	92.1	253	1	PCT-US01-19110-1007	Sequence 1007, App
23	82	92.1	253	22	US-09-880-748-1003	Sequence 1003, App
24	82	92.1	253	22	US-09-880-748-1007	Sequence 1007, App
25	82	92.1	262	22	US-09-811-737-19	Sequence 19, App1
26	82	92.1	312	14	US-09-079-029-10	Sequence 10, App1
27	80	89.9	127	22	US-09-834-366-13608	Sequence 13608, App
28	80	89.9	127	24	US-60-197-873-13608	Sequence 13608, App
29	79	88.8	92	24	US-09-834-366-13607	Sequence 13607, App
30	79	88.8	92	24	US-60-197-873-13607	Sequence 73, App1
31	75	84.3	17	13	US-08-905-825-73	Sequence 98, App1
32	75	84.3	17	15	US-09-129-026-98	Sequence 98, App1
33	75	84.3	128	22	US-09-834-366-13602	Sequence 13602, App
34	75	84.3	128	24	US-60-197-873-13602	Sequence 13602, App
35	75	84.3	250	1	PCT-US01-19110-1319	Sequence 1319, App
36	75	84.3	250	1	PCT-US01-19110-1486	Sequence 1486, App
37	75	84.3	250	22	US-09-880-748-1319	Sequence 1319, App
38	75	84.3	250	22	US-09-880-748-1486	Sequence 1486, App
39	74	83.1	254	1	PCT-US01-19110-1511	Sequence 1511, App
40	74	83.1	254	22	US-09-880-748-1511	Sequence 1511, App
41	73	82.0	17	15	US-09-129-026-110	Sequence 110, App
42	73	82.0	254	1	PCT-US01-19110-1350	Sequence 1350, App
43	73	82.0	254	22	US-09-880-748-1350	Sequence 1350, App
44	72	80.9	17	1	PCT-US00-07946-340	Sequence 340, App
45	72	80.9	17	19	US-09-534-717-340	Sequence 340, App

ALIGNMENTS

RESULT 1
PCT-US00-18152-62
Sequence 62, Application PC/TUS0018152
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
TITLE OR INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-211-228
CURRENT APPLICATION NUMBER: PCT/US00/18152
CURRENT FILING DATE: 2000-06-30
EARLIER APPLICATION NUMBER: 09/503,387
EARLIER FILING DATE: 2/14/00
EARLIER APPLICATION NUMBER: 09/454,824
EARLIER FILING DATE: 12/6/99
EARLIER APPLICATION NUMBER: 09/345,468
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 62
LENGTH: 17
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-18152-62

Query Match 100.0% Score 89; DB 1; Length 17;
Best Local Similarity 100.0% Pred. No. 6.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 NIKDGSSEKYYADSVRG 17
|||||
Db 1 NIKDGSSEKYYADSVRG 17

```
RESULT 2
US-09-610-118-62
; Sequence 62, Application US/09610118
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Petrus, M.
; APPLICANT: Valinchenker, W.
; APPLICANT: Gili, D.
; APPLICANT: Olan, M.
; APPLICANT: Kingsbury, G.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-211
; CURRENT APPLICATION NUMBER: US/09/610,118
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2/14/00
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 12/6/99
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 6/30/99
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-610-118-62

Query Match          100.0%; Score 89; DB 20; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIKODGSEKYYADSVRG 17
Db 1 NIKODGSEKYYADSVRG 17

RESULT 3
US-09-832-312-62
; Sequence 62, Application US/09832312
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-62

Query Match          100.0%; Score 89; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIKODGSEKYYADSVRG 17
Db 1 NIKODGSEKYYADSVRG 17

RESULT 4
PCT-US00-07946-664
; Sequence 664, Application PC/TUS0007946
; GENERAL INFORMATION:
; APPLICANT: Jochen, Saifeld et al.
; TITLE OF INVENTION: Human Antibodies That Bind Human IL-12 And Methods For Product
; FILE REFERENCE: BBI-093CPC
; CURRENT APPLICATION NUMBER: PCT/US00/07946
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/126,603
; EARLIER FILING DATE: March 25, 1999
; NUMBER OF SEQ ID NOS: 675
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 664
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-07946-664

Query Match          92.1%; Score 82; DB 1; Length 98;
Best Local Similarity 88.2%; Pred. No. 8.7e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODGSEKYYADSVRG 17
Db 50 NIKODGSEKYYADSVRG 66

RESULT 5
PCT-US00-07946-666
; Sequence 666, Application PC/TUS0007946
; GENERAL INFORMATION:
; APPLICANT: Jochen, Saifeld et al.
; TITLE OF INVENTION: Human Antibodies That Bind Human IL-12 And Methods For Product
; FILE REFERENCE: BBI-093CPC
; CURRENT APPLICATION NUMBER: PCT/US00/07946
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/126,603
; EARLIER FILING DATE: March 25, 1999
; NUMBER OF SEQ ID NOS: 675
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 666
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-07946-666

Query Match          92.1%; Score 82; DB 1; Length 98;
Best Local Similarity 88.2%; Pred. No. 8.7e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODGSEKYYADSVRG 17
Db 50 NIKODGSEKYYADSVRG 66

RESULT 6
US-08-124-469-32
; Sequence 32, Application US/08124469
; GENERAL INFORMATION:
; APPLICANT: RAPOPORT, Basil
; APPLICANT: MCLACHLAN, Sandra M.
; TITLE OF INVENTION: GRAVES' OPHTHALMOPATHY ASSOCIATED
; TITLE OF INVENTION: ANTIBODIES, GRAVES' OPHTHALMOPATHY ORBITAL ANTIGEN,
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR
; STREET: 1455 PENNSYLVANIA AVE., N.W.
```

CITY: WASHINGTON,
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/124,469
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 102,105,301
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-124-469-32

Query Match 92.1%; Score 82; DB 5; Length 98;
Best Local Similarity 88.2%; Pred. No. 8,7e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKDGSEKYYDSVKG 17
|||||
DB 50 NIKDGSEKYYDSVKG 66

RESULT 7
US-08-472-771-32
Sequence 32, Application US/08472771
GENERAL INFORMATION:
APPLICANT: RAPOPORT, Basil
APPLICANT: MCLACHLAN, Sandra M.
TITLE OF INVENTION: GRAVES' OPHTHALMOPATHY ASSOCIATED
TITLE OF INVENTION: ANTIBODIES, GRAVES' OPHTHALMOPATHY ORBITAL ANTIGEN,
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR
STREET: 1455 PENNSYLVANIA AVE., N.W.
CITY: WASHINGTON,
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,771
FILING DATE: Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/124,469
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 102105,301DIY
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-771-32

Query Match 92.1%; Score 82; DB 8; Length 98;
Best Local Similarity 88.2%; Pred. No. 8,7e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKDGSEKYYDSVKG 17
|||||
DB 50 NIKDGSEKYYDSVKG 66

RESULT 8
US-08-486-202-32
Sequence 32, Application US/08486202
GENERAL INFORMATION:
APPLICANT: RAPOPORT, Basil
APPLICANT: MCLACHLAN, Sandra M.
TITLE OF INVENTION: GRAVES' OPHTHALMOPATHY ASSOCIATED
TITLE OF INVENTION: ANTIBODIES, GRAVES' OPHTHALMOPATHY ORBITAL ANTIGEN,
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR
STREET: 1455 PENNSYLVANIA AVE., N.W.
CITY: WASHINGTON,
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,202
FILING DATE: Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/124,469
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 102105,301DIY
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-486-202-32

Query Match 92.1%; Score 82; DB 8; Length 98;
Best Local Similarity 88.2%; Pred. No. 8,7e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NIKDGSEKYYDSVKG 17
|||||
DB 50 NIKDGSEKYYDSVKG 66

RESULT 9
US-09-534-717-664


```

; CURRENT APPLICATION NUMBER: US/60/141,701
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 75
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-60-141-701-75

```

```

Query Match          92.1%; Score 82; DB 24; Length 125;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 NIKODGSEKYYADSVRG 17
   |||||||||
Db 51 NIKODGSEKYYVDSYKG 67

```

```

RESULT 14
US-09-811-737-3
; Sequence 3, Application US/09811737
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim Pharma KG
; TITLE OF INVENTION: Human FAP-alpha-specific antibodies
; FILE REFERENCE: 1-1129
; CURRENT APPLICATION NUMBER: US/09/811,737
; CURRENT FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-737-3

```

```

Query Match          92.1%; Score 82; DB 22; Length 134;
Best Local Similarity 88.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 NIKODGSEKYYADSVRG 17
   |||||||||
Db 50 NIKODGSEKYYVDSYKG 66

```

```

RESULT 15
PCT-US01-19110-1889
; Sequence 1889, Application PC/TUS0119110
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23PCT
; CURRENT APPLICATION NUMBER: PCT/US01/19110
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1889
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-19110-1889

```

```

Query Match          92.1%; Score 82; DB 1; Length 241;
Best Local Similarity 88.2%; Pred. No. 2.6e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 NIKODGSEKYYADSVRG 17
   |||||||||
Db 50 NIKODGSEKYYVDSYKG 66

```

```

Search completed: November 19, 2001, 08:23:14
Job time: 727 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:31 ; Search time 11.44 seconds
(without alignments)
7.915 Million cell updates/sec

Title: US-09-610-118-62

Sequence: 1 NIKODSEKRYADSVRG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 25229 seqs, 5326477 residues

Total number of hits satisfying chosen parameters: 25229

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA-New:*
1: /cgn2_6/ptodata/1/paa/PCN_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	71.9	17	5	US-09-972-656-48
2	64	71.9	225	5	US-09-453-234-102
3	64	71.9	225	5	US-09-453-234-108
4	64	71.9	225	5	US-09-453-234-110
5	64	71.9	227	5	US-09-972-656-76
6	63	70.8	179	5	US-09-782-397-2
7	63	70.8	287	5	US-09-782-397-17
8	63	70.8	304	5	US-09-782-397-14
9	62	69.7	17	5	US-09-828-708-24
10	62	69.7	101	5	US-09-828-708-10
11	61	68.5	17	5	US-09-828-708-23
12	61	68.5	102	5	US-09-972-656-126
13	60	67.4	225	5	US-09-453-234-94
14	59	66.3	225	5	US-09-453-234-56
15	59	66.3	225	5	US-09-453-234-60
16	59	66.3	225	5	US-09-453-234-92
17	59	66.3	225	5	US-09-453-234-106
18	58	65.2	225	5	US-09-453-234-100
19	54	60.7	111	1	PCF-US01-32234-7
20	50	56.2	17	5	US-09-828-708-22
21	50	56.2	112	5	US-09-828-708-8
22	50	56.2	268	1	PCF-US01-31857-1
23	50	56.2	268	5	US-09-976-118-1
24	47	52.8	17	5	US-09-972-656-45
25	47	52.8	136	5	US-09-837-306-184
26	47	52.8	223	5	US-09-972-656-86
27	47	52.8	224	5	US-09-972-656-68

28	47	52.8	225	5	US-09-453-234-68	Sequence 68, Appl
29	47	52.8	236	5	US-09-453-234-64	Sequence 64, Appl
30	47	52.8	236	5	US-09-453-234-104	Sequence 104, Appl
31	47	52.8	368	5	US-09-837-306-196	Sequence 196, Appl
32	47	52.8	16	5	US-09-972-656-52	Sequence 52, Appl
33	45	50.6	102	5	US-09-972-656-123	Sequence 123, Appl
34	45	50.6	120	5	US-09-972-656-127	Sequence 127, Appl
35	45	50.6	102	5	US-09-972-656-78	Sequence 78, Appl
36	45	50.6	463	5	US-09-976-594-721	Sequence 721, Appl
37	43	48.3	225	5	US-09-453-234-96	Sequence 96, Appl
38	42	47.2	17	5	US-09-974-449-42	Sequence 42, Appl
39	42	47.2	115	5	US-09-974-449-35	Sequence 35, Appl
40	42	47.2	218	5	US-09-974-449-35	Sequence 62, Appl
41	41	46.1	225	5	US-09-453-234-62	Sequence 125, Appl
42	40	44.9	102	5	US-09-972-656-125	Sequence 58, Appl
43	40	44.9	225	5	US-09-453-234-66	Sequence 66, Appl
44	40	44.9	225	5	US-09-453-234-66	Sequence 98, Appl
45	40	44.9	225	5	US-09-453-234-98	

ALIGNMENTS

RESULT 1
US-09-972-656-48
; Sequence 48, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-48

Query Match 71.9% Score 64: DB 5: Length 17:
Best Local Similarity 75.0% Pred. No. 3.7e-05:
Matches 12: Conservative 1: Mismatches 3: Indels 0: Gaps 0:
OY 2 IKODSEKRYADSVRG 17
DB 2 ISYDSNRYADSVRG 17
RESULT 2
US-09-453-234-102
; Sequence 102, Application US/09453234
; GENERAL INFORMATION:
; APPLICANT: Buechler, Joe
; APPLICANT: Walkler, Gunars
; APPLICANT: Gray, Jeff
; APPLICANT: Lomborg, Nils
; APPLICANT: Biosite Diagnostics, Inc.
; TITLE OF INVENTION: Human Antibodies
; FILE REFERENCE: 020015-000100S
; CURRENT APPLICATION NUMBER: US/09/453,234
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/157,415
; PRIOR FILING DATE: 1999-10-02
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 225
; TYPE: PRT

ORGANISM: Homo sapiens
OTHER INFORMATION: M2-31H
US-09-453-234-102

Query Match 71.9%; Score 64; DB 5; Length 225;
Best Local Similarity 75.0%; Pred. No. 0.00057;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKODSEKYYADSVRG 17
| ||| |||||:|
DB 51 ISYDGSNKYYADSVKG 66

RESULT 3
US-09-453-234-108
Sequence 108, Application US/09453234
GENERAL INFORMATION:
APPLICANT: Buechler, Joe
APPLICANT: Valkirs, Gunars
APPLICANT: Gray, Jeff
APPLICANT: Lonberg, Nils
APPLICANT: Biosite Diagnostics, Inc.
APPLICANT: Genpharm International
TITLE OF INVENTION: Human Antibodies
FILE REFERENCE: 020015-000110US
CURRENT APPLICATION NUMBER: US/09/453,234
CURRENT FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: US 60/157,415
PRIOR FILING DATE: 1999-10-02
NUMBER OF SEQ ID NOS: 112
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 108
LENGTH: 225
TYPE: PRT
ORGANISM: Homo sapiens
OTHER INFORMATION: M2-34H
US-09-453-234-108

Query Match 71.9%; Score 64; DB 5; Length 225;
Best Local Similarity 75.0%; Pred. No. 0.00057;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKODSEKYYADSVRG 17
| ||| |||||:|
DB 51 ISYDGSNKYYADSVKG 66

RESULT 4
US-09-453-234-110
Sequence 110, Application US/09453234
GENERAL INFORMATION:
APPLICANT: Buechler, Joe
APPLICANT: Valkirs, Gunars
APPLICANT: Gray, Jeff
APPLICANT: Lonberg, Nils
APPLICANT: Biosite Diagnostics, Inc.
APPLICANT: Genpharm International
TITLE OF INVENTION: Human Antibodies
FILE REFERENCE: 020015-000110US
CURRENT APPLICATION NUMBER: US/09/453,234
CURRENT FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: US 60/157,415
PRIOR FILING DATE: 1999-10-02
NUMBER OF SEQ ID NOS: 112
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 110
LENGTH: 225
TYPE: PRT
ORGANISM: Homo sapiens
OTHER INFORMATION: M2-35H
US-09-453-234-110

Query Match 71.9%; Score 64; DB 5; Length 225;
Best Local Similarity 75.0%; Pred. No. 0.00057;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKODSEKYYADSVRG 17
| ||| |||||:|
DB 51 ISYDGSNKYYADSVKG 66

RESULT 5
US-09-972-656-76
Sequence 76, Application US/09972656
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
APPLICANT: Tsai, Mei-Mei
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn Version 3.0
SEQ ID NO 76
LENGTH: 227
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-76

Query Match 71.9%; Score 64; DB 5; Length 227;
Best Local Similarity 75.0%; Pred. No. 0.00058;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKODSEKYYADSVRG 17
| ||| |||||:|
DB 51 ISYDGSNKYYADSVKG 66

RESULT 6
US-09-782-397-2
Sequence 2, Application US/09782397
GENERAL INFORMATION:
APPLICANT: Dan, Michael D.
Kaplan, Howard A.
TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING T FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND DETECTION OF CANCERS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Morrison & Foerster LLP
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/782,397
FILING DATE: 13-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/862,124
FILING DATE: 1997-05-22
ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.

REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 31608-20001.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 179 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-782-397-2

Query Match 70.8%; Score 63; DB 5; Length 179;
Best Local Similarity 75.0%; Pred. No. 0.00065;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
| ||| ||||| :|
Db 112 ISYDGSSTKYADSVKG 127

RESULT 7

US-09-782-397-17
Sequence 17, Application US/09782397
GENERAL INFORMATION:

APPLICANT: Dan, Michael D.
Maitl, Pradip K.
Kaplan, Howard A.

TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND DETECTION OF CANCERS

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster LLP
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/782,397
FILING DATE: 13-Feb-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/862,124
FILING DATE: 1997-05-22

ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.

REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 31608-20001.20

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:
LENGTH: 287 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-782-397-17

Query Match 70.8%; Score 63; DB 5; Length 287;
Best Local Similarity 75.0%; Pred. No. 0.0011;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
| ||| ||||| :|
Db 193 ISYDGSSTKYADSVKG 208

RESULT 8

US-09-782-397-14
Sequence 14, Application US/09782397
GENERAL INFORMATION:

APPLICANT: Dan, Michael D.
Maitl, Pradip K.
Kaplan, Howard A.

TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND DETECTION OF CANCERS

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster LLP
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/782,397
FILING DATE: 13-Feb-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/862,124
FILING DATE: 1997-05-22

ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.

REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 31608-20001.20

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 304 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-782-397-14

Query Match 70.8%; Score 63; DB 5; Length 304;
Best Local Similarity 75.0%; Pred. No. 0.0011;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
| ||| ||||| :|
Db 210 ISYDGSSTKYADSVKG 225

RESULT 9

US-09-828-708-24
Sequence 24, Application US/09828708
GENERAL INFORMATION:
APPLICANT: Ditzel, H.

RESULT 13
 US-09-453-234-94
 : Sequence 94, Application US/09453234
 : GENERAL INFORMATION:
 : APPLICANT: Buechler, Joe
 : APPLICANT: Valters, Cnars
 : APPLICANT: Gray, Jeff
 : APPLICANT: Lomborg, Nils
 : APPLICANT: Riosite Diagnostics, Inc.
 : APPLICANT: Genham International
 : TITLE OF INVENTION: Human Antibodies
 : FILE REFERENCE: 020015-00011005
 : CURRENT APPLICATION NUMBER: US/09/453, 234
 : CURRENT FILING DATE: 1999-12-01
 : PRIOR APPLICATION NUMBER: US 60/157, 415
 : PRIOR FILING DATE: 1999-10-02
 : NUMBER OF SEQ ID NOS: 112
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 94

```

; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: M2-12H
US-09-453-234-94

```

```

Query Match 67.4%; Score 60; DB 5; Length 225;
Best Local Similarity 68.8%; Pred. No. 0.0025;
Matches 11: Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 2 IKDGSSEKYYADSVRG 17
   1 11 1111111:1
Db 51 ITYDGDNKYYADSVKYG 66

```

```

RESULT 14
US-09-453-234-56
; Sequence 56, Application US/09453234
; GENERAL INFORMATION:
; APPLICANT: Buechler, Joe
; APPLICANT: Walkers, Gunars
; APPLICANT: Gray, Jeff
; APPLICANT: Lomborg, Nils
; APPLICANT: Biosite Diagnostics, Inc.
; APPLICANT: Genpharm International
; TITLE OF INVENTION: Human Antibodies
; FILE REFERENCE: 020015-000110US
; CURRENT APPLICATION NUMBER: US/09/453,234
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/157,415
; PRIOR FILING DATE: 1999-10-02
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: M1-3H
US-09-453-234-56

```

```

Query Match 66.3%; Score 59; DB 5; Length 225;
Best Local Similarity 68.8%; Pred. No. 0.0037;
Matches 11: Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 2 IKDGSSEKYYADSVRG 17
   1 11 1111111:1
Db 51 ITYDGDNKYYADSVKYG 66

```

```

RESULT 15
US-09-453-234-60
; Sequence 60, Application US/09453234
; GENERAL INFORMATION:
; APPLICANT: Buechler, Joe
; APPLICANT: Walkers, Gunars
; APPLICANT: Gray, Jeff
; APPLICANT: Lomborg, Nils
; APPLICANT: Biosite Diagnostics, Inc.
; APPLICANT: Genpharm International
; TITLE OF INVENTION: Human Antibodies
; FILE REFERENCE: 020015-000110US
; CURRENT APPLICATION NUMBER: US/09/453,234
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/157,415
; PRIOR FILING DATE: 1999-10-02
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

; OTHER INFORMATION: M1-5H
US-09-453-234-60

```

```

Query Match 66.3%; Score 59; DB 5; Length 225;
Best Local Similarity 68.8%; Pred. No. 0.0037;
Matches 11: Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 2 IKDGSSEKYYADSVRG 17
   1 11 1111111:1
Db 51 ITYDGDNKYYADSVKYG 66

```

```

Search completed: November 19, 2001, 08:23:31
Job time: 743 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:13:24 ; Search time 43.5 Seconds

(without alignments)
8,794 Million cell updates/sec

Title: US-09-610-118-62

Perfect score: 1 NIKODSEKRYADSVRG 17

Sequence: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	82	92.1	98	1	US-07-942-245-37
2	82	92.1	117	3	US-08-545-809A-95
3	69	77.5	116	3	US-08-545-809A-134
4	67	75.3	117	4	US-09-240-274-141
5	65	73.0	114	3	US-08-545-809A-124
6	65	73.0	126	1	US-08-478-039-95
7	65	73.0	126	1	US-08-478-039-95
8	64	71.9	98	1	US-08-211-202-118
9	64	71.9	116	1	US-08-211-202-118
10	64	71.9	117	3	US-08-545-809A-115
11	64	71.9	119	2	US-08-331-398A-46
12	64	71.9	119	2	US-08-331-397B-46
13	64	71.9	119	2	US-08-759-804A-46
14	64	71.9	119	4	US-09-227-693-46
15	64	71.9	120	1	US-08-211-202-135
16	64	71.9	120	1	US-07-942-245-35
17	64	71.9	123	4	US-08-983-607-38
18	64	71.9	125	1	US-08-276-852-72
19	64	71.9	125	1	US-08-276-852-72
20	64	71.9	125	1	US-08-276-852-74
21	64	71.9	125	1	US-08-276-852-75
22	64	71.9	125	1	US-08-276-852-76
23	64	71.9	125	1	US-08-276-852-77
24	64	71.9	125	1	US-08-899-575-72
25	64	71.9	125	1	US-08-899-575-73
26	64	71.9	125	1	US-08-899-575-74
27	64	71.9	125	1	US-08-899-575-75

28	64	71.9	125	1	US-08-899-575-76	Sequence 76, Appl
29	64	71.9	125	1	US-08-899-575-77	Sequence 77, Appl
30	64	71.9	125	1	US-08-899-575-72	Sequence 72, Appl
31	64	71.9	125	1	US-08-899-575-73	Sequence 73, Appl
32	64	71.9	125	1	US-08-899-575-75	Sequence 74, Appl
33	64	71.9	125	1	US-08-899-575-75	Sequence 75, Appl
34	64	71.9	125	1	US-08-899-575-76	Sequence 76, Appl
35	64	71.9	125	1	US-08-899-575-77	Sequence 77, Appl
36	64	71.9	125	5	PCT-US95-08743-72	Sequence 72, Appl
37	64	71.9	125	5	PCT-US95-08743-73	Sequence 73, Appl
38	64	71.9	125	5	PCT-US95-08743-74	Sequence 74, Appl
39	64	71.9	125	5	PCT-US95-08743-75	Sequence 75, Appl
40	64	71.9	125	5	PCT-US95-08743-76	Sequence 76, Appl
41	64	71.9	125	5	PCT-US95-08743-77	Sequence 77, Appl
42	64	71.9	128	1	US-08-478-039-96	Sequence 96, Appl
43	64	71.9	128	1	US-08-476-349A-96	Sequence 96, Appl
44	63	70.8	179	4	US-08-862-124-2	Sequence 2, Appl
45	63	70.8	287	4	US-08-862-124-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-37
; Sequence 37, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Sughrie, Mlon, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: HP 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELE: 6491103
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-942-245-37

Query Match 92.1%; Score 82; DB 1; Length 98;
Best Local Similarity 82.2%; Pred. No. 3e-07;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODSEKRYADSVRG 17
Db 50 NIKODSEKRYADSVRG 66

```

RESULT 2
US-08-545-809A-95
; Sequence 95, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Matsuda, Fumihiko
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,809A
; FILING DATE: 27-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP93/00603
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 06501/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-342-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-545-809A-95

Query Match 92.1%; Score 82; DB 3; Length 117;
Best Local Similarity 88.2%; Pred. No. 3,7e-07;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NIKODGSEKYYADSVRG 17
| | | | | | | | | | | | | | | | | | | | | |
Db 69 NIKODGSEKYYDVSKG 85

RESULT 3
US-08-545-809A-134
; Sequence 134, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Matsuda, Fumihiko
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:

```

```

1  MEDIUM TYPE: Diskette
2  COMPUTER: IBM Compatible
3  OPERATING SYSTEM: Windows95
4  SOFTWARE: FASTEDQ for Windows Version 2.0.C
5  CURRENT APPLICATION DATA:
6  APPLICATION NUMBER: US/08/545,809A
7  FILING DATE: 27-MAR-1996
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: PCT/JP93/00603
10 FILING DATE: 10-MAY-1993
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Freeman, John W.
13 REGISTRATION NUMBER: 29,066
14 REFERENCE/DOCKET NUMBER: 06501/004001
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 617-542-5070
17 TELEFAX: 617-542-8906
18 TELEX: 200154
19 INFORMATION FOR SEQ. ID NO.: 134:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 116 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25
26 US-08-545-809A-134

```

Query Match 77.5%; Score 69; DB 3; Length 116;
Best Local Similarity 76.5%; Pred. NO. 6.7e-05;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
QY      1 NIKQDGESEKYADSVRG 17
          :|| ||||| |||:|
Db      68 DIKCDGESEKYYVDSVKG 84
```

RESULT 4
 US-09-240-274-141
 : Sequence 141, Application US/09240274
 : Patent No. 6253455
 : GENERAL INFORMATION:
 : APPLICANT: Siegel, Donald L.
 : TITLE OF INVENTION: Rh(d)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
 : TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
 : FILE REFERENCE: 09596-4202
 : CURRENT APPLICATION NUMBER: US/09/240, 274
 : CURRENT FILING DATE: 1999-01-29
 : EARLIER APPLICATION NUMBER: 60/081,380
 : EARLIER FILING DATE: 1998-04-10
 : EARLIER APPLICATION NUMBER: 60/026,550
 : EARLIER FILING DATE: 1996-10-11
 : NUMBER OF SEQ ID NOS: 224
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 141
 : LENGTH: 117
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : FEATURE:
 : OTHER INFORMATION: anti-Rh(D) antibody clone SH17
 : US-09-240-274-141

```

Qy      1 NIKODSEKRYADSVNC 17
      |||||||:| ||| |
Db      50 NIKODSEKRYADSVNC 66
      |||||||:| ||| |

Query Match      75.3%; Score 67; DB 4; Length 117;
Best Local Similarity 76.5%; Pred. No. 0.00015;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0.

```

RESULT 5
US-08-545-809A-124

;; FILING DATE: 25-JAN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/912,292
;; FILING DATE: 10-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/856,281
;; FILING DATE: 23-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/735,064
;; FILING DATE: 25-JUL-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Teskin Esq., Robin L.
;; REGISTRATION NUMBER: 35,030
;; REFERENCE/DOCKET NUMBER: 012712-161
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-836-6620
;; TELEFAX: 703-836-2021
;; INFORMATION FOR SEQ ID NO: 95:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 126 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; POSITION IN GENOME:
;; CHROMOSOME/SEGMENT: RF SJ1
US-08-476-349A-95

Query Match 73.0%; Score 65; DB 1; Length 126;
Best Local Similarity 75.0%; Pred. No. 0.00036;
Matches 12: Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IKODSEKYYADSVRG 17
+ + + + +
Db 51 ISDDGSKYYADSVKG 66

RESULT 8
US-08-211-202-118
;; Sequence 118, Application US/08211202
;; Patent No. 5565332
;; GENERAL INFORMATION:
;; APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
;; APPLICANT: BAIER, Michael
;; APPLICANT: JESPERS, Laurent Stephane Anne Therese
;; APPLICANT: WINTER, Gregory Paul
;; TITLE OF INVENTION: Production of chimeric antibodies - a
;; TITLE OF INVENTION: combinatorial approach
;; NUMBER OF SEQUENCES: 144
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
;; ADDRESSEE: Borun
;; STREET: 6300 Sears Tower, 233 South Wacker Drive
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/211,202
;; FILING DATE: 23-SEP-1992
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9120252.3
;; FILING DATE: 23-SEP-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9120377.8
;; FILING DATE: 25-SEP-1991

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9206318.9
;; FILING DATE: 24-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9206372.6
;; FILING DATE: 24-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/GB92/00883
;; FILING DATE: 15-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: David W. Clough
;; REGISTRATION NUMBER: 36,107
;; REFERENCE/DOCKET NUMBER: 28111/31960
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-474-6300
;; TELEFAX: 312-474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 118:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 98 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-211-202-118

Query Match 71.9%; Score 64; DB 1; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.00041;
Matches 12: Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IKODSEKYYADSVRG 17
+ + + + +
Db 51 ISYDGNKYYADSVKG 66

RESULT 9
US-08-211-202-141
;; Sequence 141, Application US/08211202
;; Patent No. 5565332
;; GENERAL INFORMATION:
;; APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
;; APPLICANT: BAIER, Michael
;; APPLICANT: JESPERS, Laurent Stephane Anne Therese
;; APPLICANT: WINTER, Gregory Paul
;; TITLE OF INVENTION: Production of chimeric antibodies - a
;; TITLE OF INVENTION: combinatorial approach
;; NUMBER OF SEQUENCES: 144
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
;; ADDRESSEE: Borun
;; STREET: 6300 Sears Tower, 233 South Wacker Drive
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/211,202
;; FILING DATE: 23-SEP-1992
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9120252.3
;; FILING DATE: 23-SEP-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9120377.8
;; FILING DATE: 25-SEP-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9206318.9
;; FILING DATE: 24-MAR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 141:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-211-202-141

Query Match 71.9%; Score 64; DB 1; Length 116;
Best Local Similarity 75.0%; Pred. No. 0.0005;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
| ||| ||||| ||| :1
DB 51 ISYDGSNKYYADSVRG 66

RESULT 10
US-08-545-809A-115
Sequence 115, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-545-809A-115

Query Match 71.9%; Score 64; DB 3; Length 117;
Best Local Similarity 75.0%; Pred. No. 0.0005;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKDGSSEKYYADSVRG 17
| ||| ||||| ||| :1
DB 70 ISYDGSNKYYADSVRG 85

RESULT 11
US-08-331-398A-46
Sequence 46, Application US/08331398A
Patent No. 5608039
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Willingham, Mark
APPLICANT: Fitzgerald, David
APPLICANT: Brinkmann, Ulrich
APPLICANT: Pai, Lee
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins
TITLE OF INVENTION: and Their Uses (as amended)
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Plaza
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,398A
FILING DATE: 28-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-126110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..119
OTHER INFORMATION:
OTHER INFORMATION: 56p1/CL Variable Heavy chain (V-H)*
US-08-331-398A-46

Query Match 71.9%; Score 64; DB 1; Length 119;
Best Local Similarity 75.0%; Pred. No. 0.00051;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKODSEKYYADSVRG 17
| ||| |||||:1
Db 51 ISYDGSNKYYADSVKG 66

RESULT 12
US-08-331-397B-46
; Sequence 46, Application US/08331397B
; Patent No. 5981726
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Benhar, Itai
; TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-
; TITLE OF INVENTION: Specific Antibody Fragments, Fusion Proteins, and Uses
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Stewart Street Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,397B
; FILING DATE: 28-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/596,289
; FILING DATE: 12-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-126120US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: protein
; LOCATION: 1..119
; OTHER INFORMATION: /note="Human fetal immunoglobulin
; OTHER INFORMATION: 56p1/CL Variable Heavy chain (V-H)"
US-08-331-397B-46

Query Match 71.9%; Score 64; DB 2; Length 119;
Best Local Similarity 75.0%; Pred. No. 0.00051;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKODSEKYYADSVRG 17
| ||| |||||:1
Db 51 ISYDGSNKYYADSVKG 66

RESULT 13
US-08-759-804A-46
; Sequence 46, Application US/08759804A

; Patent No. 5990296
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Willingham, Mark
; APPLICANT: Filzgerald, David J.
; APPLICANT: Brinkmann, Ulrich
; APPLICANT: Pal, Lee
; TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
; TITLE OF INVENTION: Fusion Proteins, and Uses Thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,804A
; FILING DATE: 03-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,398
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/596,289
; FILING DATE: 12-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-126140US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: protein
; LOCATION: 1..119
; OTHER INFORMATION: /note="Human fetal immunoglobulin
; OTHER INFORMATION: 56p1/CL Variable Heavy chain (V-H)"
US-08-759-804A-46

Query Match 71.9%; Score 64; DB 2; Length 119;
Best Local Similarity 75.0%; Pred. No. 0.00051;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 IKODSEKYYADSVRG 17
| ||| |||||:1
Db 51 ISYDGSNKYYADSVKG 66

RESULT 14
US-09-227-693-46
; Sequence 46, Application US/09227693
; Patent No. 6287562
; GENERAL INFORMATION:
; APPLICANT: PASTAN, Ira
; APPLICANT: BENHAR, Itai

APPLICANT: PADLAN, Eduardo A.
APPLICANT: JUNG, Sun-Hee
APPLICANT: LEE, Byungkook
TITLE OF INVENTION: HUMANIZED TUMOR-SPECIFIC ANTIBODY
TITLE OF INVENTION: FRAGMENTS, FUSION PROTEINS, AND USES THEREOF
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/227,693
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/331,396
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/767,331
FILING DATE: 30-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/596,289
FILING DATE: 12-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 15280-126-1-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..119
OTHER INFORMATION: /note="Human fetal immunoglobulin
OTHER INFORMATION: 56pI/CL VH region"
US-09-227-693-46
Query Match 71.9%; Score 64; DB 4; Length 119;
Best Local Similarity 75.0%; Pred. No. 0.00051;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 2 IKDGSERYADSVRG 17
Db 51 ISYDGSNKRYADSVKG 66
RESULT 15
US-08-211-202-135
Sequence 135, Application US/08211202
Patent No. 5565332
GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
APPLICANT: BAIER, Michael
APPLICANT: JESPEERS, Laurent Stephane Anne Therese
APPLICANT: WINTER, Gregory Paul
TITLE OF INVENTION: Production of chimeric antibodies - a
TITLE OF INVENTION: combinatorial approach
NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:
ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120252.3
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120377.8
FILING DATE: 25-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 135:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-211-202-135

Query Match 71.9%; Score 64; DB 1; Length 120;
Best Local Similarity 75.0%; Pred. No. 0.00051;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 2 IKDGSERYADSVRG 17
Db 51 ISYDGSNKRYADSVKG 66

Search completed: November 19, 2001, 08:13:24
Job time: 137 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:14:20 ; Search time 48.99 Seconds
(without alignments)
21.769 Million cell updates/sec

Title: US-09-610-118-63

Perfect score: 80

Sequence: 1 DKWEAYITPGAFDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR.68: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	57.5	484	2	T38463
2	45	56.2	765	2	S76795
3	44	55.0	175	2	T00186
4	44	55.0	682	2	T50840
5	44	55.0	683	2	T50839
6	44	55.0	867	1	C64785
7	44	55.0	869	2	E85553
8	43	53.8	653	2	F85620
9	43	53.8	866	1	C64834
10	42	52.5	108	2	PH1015
11	41	51.2	312	2	E71624
12	41	51.2	378	2	T50781
13	41	51.2	380	2	A75328
14	41	51.2	411	2	S07472
15	41	51.2	422	2	T06388
16	41	51.2	684	2	H82296
17	41	51.2	796	2	T23238
18	40	50.0	721	2	C71014
19	40	50.0	3158	2	T17483
20	39	48.8	148	2	F71207
21	39	48.8	179	2	D71957
22	39	48.8	264	2	F70828
23	39	48.8	367	2	A86029
24	39	48.8	434	2	T47748
25	39	48.8	495	2	G83325
26	39	48.8	592	2	PS0197
27	39	48.8	596	1	TJ0345
28	39	48.8	766	2	T01900
29	39	48.8	998	2	T04842

30	39	48.8	1041	2	E70760	probable lies prot
31	38	47.5	71	2	T43545	DNA-directed RNA p
32	38	47.5	109	2	D70701	hypothetical prote
33	38	47.5	133	2	PC1155	ig heavy chain pre
34	38	47.5	204	2	T04423	probable alpha-gal
35	38	47.5	257	2	H84597	hypothetical prote
36	38	47.5	288	2	D82892	conserved hypotet
37	38	47.5	328	2	B84545	hypothetical prote
38	38	47.5	349	2	C72630	probable threonine
39	38	47.5	422	2	D84403	dihydroxotase [lm
40	38	47.5	448	2	B75180	antibiotic/peptide
41	38	47.5	482	2	JC7332	endoglycosylceram
42	38	47.5	656	2	T28058	hypothetical prote
43	38	47.5	659	2	C85057	probable receptor
44	38	47.5	756	2	S67433	hypothetical prote
45	38	47.5	1218	2	S71376	glutamate receptor

ALIGNMENTS

RESULT 1
T38463
probable aminotransferase (EC 2.6.1.-) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T38463
R:Harris, D.; McDonald, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A:Reference number: 221794
A:Accession: T38463
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-484 <HAR>
A:Cross-references: EMBL:269368; PIDN:CAA93294.1; GSPDB:GN00066; SPDB:SPAC27FL.05c
A:Experimental source: strain 972h; cosmid c27f1
C:Genetics:
A:Gene: SPDB:SPAC27FL.05c
A:Map position: 1
C:Superfamily: beta-alanine--pyruvate transaminase
C:Keywords: aminotransferase

Query Match 57.5%; Score 46; DB 2; Length 484;
Best Local Similarity 50.0%; Pred. No. 3.6;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFDV 14
:11:1:1111
DB 204 EKWQKYSPPFVDV 217

RESULT 2
S76795
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.

A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76795
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K.; Okumura, S.; Shimpo, S.; Takuchih, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.

A:Reference number: S74322; MUID:97061201

A:Accession: S76795
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-765 <KAN>

A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BA18707.1; PID:d101
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

A:Gene: sfmd
C:Superfamily: outer membrane usher protein fimd

Query Match 55.0%; Score 44; DB 2; Length 869;
Best Local Similarity 41.7%; Pred. No. 15;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 3 WEAYITPGAFDV 14
DB 309 YQTVSPGAFEI 320

RESULT 8
F85620
partial fimbrial usher protein [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: F85620
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206531
A:Accession: F85620
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-653 <STO>
A:Cross-references: GB:AE005174; NID:q12514119; PIDN:ANG55426.1; GSPDB:GN00145; OMGP:Z12
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1289

Query Match 53.8%; Score 43; DB 2; Length 653;
Best Local Similarity 41.7%; Pred. No. 17;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 3 WEAYITPGAFDV 14
DB 95 YQTVSPGAFEI 106

RESULT 9
C64834
probable outer membrane usher protein ycbS - Escherichia coli
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C:Accession: C64834
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: C64834
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-866 <BLAT>
A:Cross-references: GB:AE000196; GB:U00096; NID:q1787169; PIDN:AMC74026.1; PID:q1787172;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ycbS
C:Superfamily: outer membrane usher protein fimd
C:Keywords: ATP; fimbria; membrane protein; nucleotide binding; P-loop; transport protei
F:15-31/Domain: transmembrane #status predicted <TM>
F:642-649/Region: nucleotide-binding motif A (P-loop)

Query Match 53.8%; Score 43; DB 1; Length 866;
Best Local Similarity 41.7%; Pred. No. 23;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 3 WEAYITPGAFDV 14

DB 308 YQTVSPGAFEI 319

RESULT 10
PH1015
Ig heavy chain V region (clone 111.55) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PH1015
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marlon, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0971; MUID:92381444
A:Accession: PH1015
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-108 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region: immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 52.5%; Score 42; DB 2; Length 108;
Best Local Similarity 46.2%; Pred. No. 3.4;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 KWEAYITPGAFDV 14
DB 46 EWVAYISSGSFNI 58

RESULT 11
E71624
rifin PRB0055c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 02-Mar-2001
C:Accession: E71624
R:Gardner, M.J.; Tetelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743
A:Accession: E71624
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-312 <GAR>
A:Cross-references: GB:AE001368; GB:AE001362; NID:q3845079; PIDN:AC71801.1; PID:q384
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PRB0055c
C:Superfamily: Plasmodium falciparum rifin PRB1005w

Query Match 51.2%; Score 41; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 KWEAYITPGAFD 13
DB 177 KWSAMVTPPTYD 188

RESULT 12
T50781
alpha-galactosidase (EC 3.2.1.22) [imported] - coffee
C:Species: Coffea arabica (coffee)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 18-Aug-2000
C:Accession: T50781
R:Zhu, A.; Goldstein, J.
Gene 140, 227-231, 1994
A:Title: Cloning and functional expression of a cDNA encoding coffee bean alpha-galac

A:Reference number: Z25235; MUID:94193002
 A:Accession: T50781
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-378 <RHU>
 A:Cross-references: EMBL:L27992; PIDN:AAA3022.1
 C:Superfamily: alpha-galactosidase
 C:Keywords: glycosidase; hydrolase

Query Match

Best Local Similarity 51.2%; Score 41; DB 2; Length 378;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
 |||:| ||:

Db 216 DKWASYAGCGGN 228

RESULT 13

A75328

PhoH-related protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: A75328

R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; Ma

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896

A:Accession: A75328

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <WHI>

A:Cross-references: GB:AE002038; GB:AE000513; NID:96459790; PIDN:AAF11540.1; PID:9645977

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1988

A:Map position: 1

C:Superfamily: phosphate starvation-induced protein

Query Match

Best Local Similarity 51.2%; Score 41; DB 2; Length 380;
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 14
 :|:|:|:|:|:

Db 243 EKFEAYITSGVIEI 256

RESULT 14

S07472

alpha-galactosidase (EC 3.2.1.22) precursor - guar

C:Species: *Cyamopsis tetragonoloba* (guar, cluster bean)

C>Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 22-Jun-1999

C:Accession: S07472

R.Overbeek, N.; Fellinger, A.J.; Toonen, M.Y.; van Wassenaar, D.; Verrips, C.T.

Plant Mol. Biol. 13, 541-550, 1989

A:Title: Cloning and nucleotide sequence of the alpha-galactosidase cDNA from *Cyamopsis*

A:Accession: S07472

A:Molecule type: mRNA

A:Residues: 1-411 <OVE>

A:Cross-references: EMBL:X14619; NID:q18291; PIDN:CAA2772.1; PID:q18292

A>Note: the authors translated the codon ATT for residue 20 as Asn, TAT for residue 140

C:Superfamily: alpha-galactosidase

C:Keywords: glycosidase; hydrolase

Query Match

Best Local Similarity 51.2%; Score 41; DB 2; Length 411;
 46.2%; Pred. No. 22;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
 |||:| ||:

Db 248 DKWASYAGCGGN 260

RESULT 15

T06388

alpha-galactosidase (EC 3.2.1.22) - soybean

C:Species: *Glycine max* (soybean)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Jun-1999

C:Accession: T06388

R.Davis, M.O.; Walker, J.C.; Smith, D.

submitted to the EMBL Data Library, August 1994

A:Description: Cloning and expression of a soybean alpha galactosidase gene.

A:Reference number: Z15645

A:Accession: T06388

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-422 <DAV>

A:Cross-references: EMBL:U12926; NID:9927574; PIDN:AAV3963.1; PID:9927575

A:Experimental source: strain Williams

C:Function:

A:Description: catalyzes hydrolysis of melibiose into galactose and glucose

C:Superfamily: alpha-galactosidase

C:Keywords: glycosidase; hydrolase

Query Match

Best Local Similarity 51.2%; Score 41; DB 2; Length 422;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
 |||:| ||:

Db 260 DKWASYAGCGGN 272

Search completed: November 19, 2001, 08:14:21
 Job time: 194 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:34 : Search time 27.32 seconds

(Without alignments)
17.554 Million cell updates/sec

Title: US-09-610-118-63

Perfect score: 80
Sequence: 1 DKWEATYFGADV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	46	57.5	YAV5_SCHPO	Q10174 schizosacch
2	44	55.0	SFMD_ECOLI	P77468 escherichia
3	43	53.8	YCBS_ECOLI	P75857 escherichia
4	41	51.2	AGAL_COFAR	Q42656 coffea arab
5	41	51.2	AGAL_CYATE	P14749 cyamopsis t
6	41	51.2	YRS7_CAEEL	Q10003 caenorhabdi
7	40	50.0	FIMD_SALTY	P37924 salmoneilla
8	39	48.8	PHLC_BACCE	P05598 bacillus ce
9	39	48.8	PHLD_BACCE	P33376 bacillus ce
10	39	48.8	VP1_BPCHP	P19192 bacterioph
11	39	48.8	SVI_MYCTU	Q10765 mycobacteri
12	38	47.5	YAV5_SCHPO	Q13877 schizosacch
13	38	47.5	Y030_MYCTU	P71600 mycobacteri
14	38	47.5	RS3A_SULSO	O9UD4 sulfolobus
15	38	47.5	SSRA_ARATH	P42434 arabidopsis
16	38	47.5	YD34_SCHPO	Q10268 schizosacch
17	37.5	46.9	RRPO_PEAAY	P29154 pea enatinn
18	37	46.2	RRPX_YEAST	P22139 saccharomyc
19	37	46.2	MTRD_METKA	O32864 methanopyru
20	37	46.2	XINA_THEDU	P23360 thermascus
21	37	46.2	Y030_NPVAC	P41340 autographa
22	37	46.2	FTPL_BOMMO	P49687 bombyx mori
23	37	46.2	TYRO_PODAN	O92396 podospora a
24	37	46.2	GLGB_HUMAN	O04446 homo sapien
25	37	46.2	RRPO_NEUTN	P33541 neurospora
26	37	46.2	CCAG_RAT	O54898 rattus norv
27	37	46.2	CCAG_HUMAN	O43497 homo sapien
28	37	46.2	TENX_HUMAN	P22105 homo sapien
29	36	45.0	SCAB_STRHA	O05362 streptomyce
30	36	45.0	SPAO_SALTI	O56022 salmoneilla
31	36	45.0	CARA_SULSO	O59968 sulfolobus
32	36	45.0	MANB_BACSM	P16699 bacillus sp
33	36	45.0	ACSA_BACSU	P39062 bacillus su

34	36	45.0	584	1	PMEL_CITSI	O04886 citrus sine
35	36	45.0	666	1	POL_FMYD	P09523 figwort mos
36	36	45.0	674	1	POL_CAMVD	P03556 cauliflower
37	36	45.0	679	1	POL_CAMYC	P03555 cauliflower
38	36	45.0	679	1	POL_CAMVE	O02964 cauliflower
39	36	45.0	679	1	POL_CAMVS	P03554 cauliflower
40	36	45.0	680	1	POL_CAMVN	O00962 cauliflower
41	36	45.0	997	1	Y414_MYCPN	P75183 mycoplasma
42	36	45.0	1024	1	Y075_MYGCE	P47321 mycoplasma
43	36	45.0	1106	1	ACLY_CAEEL	P53585 caenorhabdi
44	36	45.0	1536	1	SIN3_YEAST	P22579 saccharomyc
45	36	45.0	4367	1	DYHC_NEUCR	P45443 neurospora

ALIGNMENTS

RESULT 1
YAV5_SCHPO STANDARD: PRT: 484 AA.

AC Q10174;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROBABLE AMINOTRANSFERASE C27F1.05C (EC 2.6.1.-).
GN SPAC27F1.05C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID:4896;
NM [1]
NM SEQUENCE FROM N.A.
RC STRAIN-972;
RA Harris D., McDonald S., Barrell B.G., Rejandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/Genbank/DDBJ databases.
CC -I- COFACTOR: PYRIDOXAL PHOSPHATE (POTENTIAL).
CC -I- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb.slb.ch).

CC EMBL: 269368; CAA93294.1; -.
CC Interpro: IPR000954; -.
DR Pfam: PF00202; amlnotran_3; 1.
DR PROSITE: PS00600; AA-TRANSFER-CLASS_3; 1.
KW Hypothetical protein; Transferase; Amino transferase;
KW Pyridoxal phosphate
FT BINDING 305
FT SEQUENCE 484 AA; 53190 MW; 5DB9A13BA33C497 CRC64;

Query Match 57.5%; Score 46; DB 1; Length 484;
Best Local Similarity 50.0%; Pred. No. 2;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
DB 204 EKWOKYSPAPFDV 217

RESULT 2
SFMD_ECOLI STANDARD: PRT: 867 AA.
AC P77468;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

Query Match	55.0%	Score 44	DB 1	Length 867
Best Local Similarity	41.7%	Pred. NO. 7.7		
Matches	5	Conservative	7	Mismatches
			0	Indels
			0	Gaps
			0	
Qy	3	WEAYITPGAFDV	14	
		::: :::		
Db	309	YOSYVSPGAFEI	320	
RESULT	3			
YCBS_ECOLI				
ID	YCBS_ECOLI	STANDARD:	PRT:	866 AA.
AC	P75857;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	HYPOTHETICAL OUTER MEMBRANE USHER PROTEIN IN PEPR-PYRD INTERGENIC			
DE	REGION PRECURSOR.			
GN	YCBS.			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			

Query Match	41.7%	Score 43	DB 1	Length 866
Best Local Similarity	41.7%	Pred. No. 11		
Matches	5	Conservative	6	Mismatches 1
				Indels 0
				Gaps 0
Qy	3	WEATYTPGAFDV	14	
		::: :::		
Db	308	YQTVSPGAFEI	319	
RESULT 4				
AGAL_COFAR				
ID	AGAL_COFAR	STANDARD	PRT	378 AA.
AC	Q42636			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	ALPHA-GALACTOSIDASE PRECURSOR (EC 3.2.1.22) (MELIBIASE) (ALPHA-D-			
DE	GALACTOSIDE GALACTOHYDROLASE).			
OS	Coffea arabica (Coffee).			
OC	Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta;			
CC	Magnoliophyta: eudicotyledons: core eudicots: Asterales: euasterids I;			

```

OC Gentianales; Rubiaceae; Coffee.
RX NCBI_TaxID=13443;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-34; 215-231 AND 373-378.
RC TISSUE=Seed;
RA MEDLINE=94193002; PubMed=8144030;
RT Zhu A., Goldstein J.;
RT "Cloning and functional expression of a cDNA encoding coffee bean
  alpha-galactosidase.";
RL Gene 140:227-231(1994).
CC -1- FUNCTION: PREFERENTIALLY CLEAVES ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDE
  LINKAGES. INVOLVED IN THE HYDROLYSIS OF THE GALACTOMANNAN, IT
  SPLITS ALPHA-LINKED GALACTOSE MOIETIES. IT IS PARTICULARLY
  SUITABLE FOR THE HYDROLYSIS OF GUAR GUM TO A GUM WITH IMPROVED
  GELLING PROPERTIES. CAN CLEAVE TERMINAL ALPHA-1,3-LINKED GALACTOSE
  RESIDUES RESPONSIBLE FOR BLOOD GROUP B SPECIFICITY FROM THE
  SURFACE OF ERYTHROCYTES THEREBY CONVERTING THESE CELLS
  SEROLOGICALLY TO GROUP O.
CC -1- CATALYTIC ACTIVITY: MELIBIOSE + H(2)O = GALACTOSE + GLUCOSE.
CC -1- SIMILARITY: BELONGS TO FAMILY 27 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
CC EMBL: L27992; AAA3022.1; .
CC DR InterPro: IPR000111; .
CC DR InterPro: IPR002241; .
CC DR InterPro: IPR002287; .
CC DR Pfam: PF02065; Melibiase; 2.
CC DR PRINTS: PR00740; GLHYDRASE27.
CC DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
CC KW Hydrolase; Glycosidase; Signal.
CC FT SIGNAL 1 15
CC FT CHAIN 16 378 ALPHA-GALACTOSIDASE.
CC FT ACT_SITE 287 287 POTENTIAL.
CC FT SEQUENCE 378 AA; 41310 MW; 9FC7610BFD760AE3 CRC64;
CC SQ

```

Query Match 51.2%; Score 41; DB 1; Length 378;
 Best Local Similarity 46.2%; Pred. NO. 11;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 DKWEAYITPGAFD 13
    111:1 11::
DB 216 DKWASYAGPGWN 228

```

RESULT 5

AGAL_CYATE STANDARD; PRT; 411 AA.

AC P14749;
 DT 01-APR-1990 (rel. 14; Created)
 DT 01-APR-1990 (rel. 14; Last sequence update)
 DT 01-NOV-1997 (rel. 35; Last annotation update)
 DE ALPHA-GALACTOSIDASE PRECURSOR (EC 3.2.1.22) (MELIBIASE) (ALPHA-D-
 GALACTOSIDE GALACTOHYDROLASE).
 OS Cynamopsis tetragonoloba (Guar) (Cluster bean).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fabales; Fabaceae; Papilionoideae; Cynampsis.
 OX NCBI_TaxID=3832;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Alleurone;
 RX MEDLINE=91370836; PubMed=2577496;
 RA Overbeke N., Fellinger A.J., Toonen M.Y., van Wassenaar D.,
 RA Verrips C.T.;
 RT "Cloning and nucleotide sequence of the alpha-galactosidase cDNA from

```

RT Cynamopsis tetragonoloba (guar).";
RL Plant Mol. Biol. 13:541-550(1989).
RN [2]
RP SEQUENCE OF 48-57 AND 172-178.
RC TISSUE=Seed;
RA Hughes S.G., Overbeke N., Robinson S., Pollock K., Smeets F.L.M.;
RT "Messenger RNA from isolated aleurone cells directs the synthesis of
  an alpha-galactosidase found in the endosperm during germination of
  guar (Cynamopsis tetragonoloba) seed.";
RL Plant Mol. Biol. 11:783-789(1988).
CC -1- FUNCTION: INVOLVED IN THE HYDROLYSIS OF THE GALACTOMANNAN, IT
  SPLITS ALPHA-LINKED GALACTOSE MOIETIES. IT IS PARTICULARLY
  SUITABLE FOR THE HYDROLYSIS OF GUAR GUM TO A GUM WITH IMPROVED
  GELLING PROPERTIES. PREFERENTIALLY CLEAVES ALPHA-1,6 GLYCOSIDE
  LINKAGES.
CC -1- CATALYTIC ACTIVITY: MELIBIOSE + H(2)O = GALACTOSE + GLUCOSE.
CC -1- SIMILARITY: BELONGS TO FAMILY 27 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X14619; CAA32772.1; .
CC DR PIR: S07472; S07472.
CC DR InterPro: IPR000111; .
CC DR InterPro: IPR002241; .
CC DR InterPro: IPR002287; .
CC DR Pfam: PF02065; Melibiase; 2.
CC DR PRINTS: PR00740; GLHYDRASE27.
CC DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
CC KW Hydrolase; Glycosidase; Signal; Glycoprotein.
CC FT SIGNAL 1 24
CC FT PROPEP 25 47
CC FT CHAIN 48 411
CC FT CARBOHYD 32 32 N-LINKED (GLCNAc . . .) (POTENTIAL).
CC FT CARBOHYD 145 145 N-LINKED (GLCNAc . . .) (POTENTIAL).
CC FT CARBOHYD 352 352 N-LINKED (GLCNAc . . .) (POTENTIAL).
CC FT ACT_SITE 319 319 POTENTIAL.
CC FT SEQUENCE 411 AA; 45135 MW; 5B171585D1AB1E CRC64;
CC SQ

```

Query Match 51.2%; Score 41; DB 1; Length 411;
 Best Local Similarity 46.2%; Pred. NO. 12;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 DKWEAYITPGAFD 13
    111:1 11::
DB 248 DKWASYAGPGWN 260

```

RESULT 6

YRS7_CAEEL STANDARD; PRT; 796 AA.

AC Q10003;
 DT 01-OCT-1996 (rel. 34; Created)
 DT 01-OCT-1996 (rel. 34; Last sequence update)
 DT 30-MAY-2000 (rel. 39; Last annotation update)
 DE HYPOTHETICAL 90.8 KDA PROTEIN T05H10.7 IN CHROMOSOME II.
 GN T05H10.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Lightning J., Thomas K.;
 RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z47811; CAAB7788.1; JOINED.
DR EMBL: Z47812; CAAB7788.1; JOINED.
DR EMBL: Z47812; CAAB7788.1; JOINED.
DR Wormpep: T05H10.7; CE03637.
KW Hypothetical protein.
FT DOMAIN 38 42 POLY-GLU.
FT DOMAIN 67 70 POLY-GLU.
FT DOMAIN 524 527 POLY-GLU.
SQ SEQUENCE 796 AA; 90831 MW; 7BDF8E0AD2AA9F1 CRC64;

Query Match 51.2%; Score 41; DB 1; Length 796;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DKWEAYTP 9
Db 163 DKWEAYTP 171

RESULT 7
FIND_SALTY STANDARD: PRT; 870 AA.
AC P37924;
DT 01-OCT-1994 (rel. 30, Created)
DT 01-OCT-1994 (rel. 30, Last sequence update)
DT 01-NOV-1995 (rel. 32, Last annotation update)
DE OUTER MEMBRANE USHER PROTEIN F1MD PRECURSOR.
GN F1MD.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RA Swenson D.L., Clegg S.;
RL Submitted (JUN-1993) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF F1MA FIMBRIAL
CC SUBUNITS ACROSS THE OUTER MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L19338; AAN75419.1; -
DR StyGene; SG10503; f1md.
DR InterPro: IPR000015; -
DR Pfam: PF00577; Usher; 1.
DR PROSITE: PS01151; FIMBRIAL_USHER; 1.
KW Outer membrane; Transmembrane; Fimbria; Transport; Signal.
FT SIGNAL 1 7 POTENTIAL.
FT CHAIN 1 870 OUTER MEMBRANE USHER PROTEIN F1MD.
FT DISULFID 843 865 POTENTIAL.
SQ SEQUENCE 870 AA; 95189 MW; 034A5FB2CF689BAF CRC64;

Query Match 50.0%; Score 40; DB 1; Length 870;

Best Local Similarity 41.7%; Pred. No. 37;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 3 WEAYTPGAFDV 14
Db 310 YOSYVPGAFAT 321

RESULT 8
PHLC_BACCE STANDARD: PRT; 283 AA.
ID PHLC_BACCE
AC P09598;
DT 01-MAR-1989 (rel. 10, Created)
DT 01-MAR-1989 (rel. 10, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last annotation update)
DE PHOSPHOLIPASE C PRECURSOR (EC 3.1.4.3) (PLC) (PHOSPHATIDYLCHOLINE
DE CHOLINEPHOSPHOHYDROLASE) (CEROLYSIN A).
GN PLC.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SE-1;
RX MEDLINE=88313678; PubMed=3137122;
RA Johansen T., Holm T., Guddal P.H., Sletten K., Haugli F.B., Little C.;
RT "Cloning and sequencing of the gene encoding the phosphatidylcholine-
RT preferring phospholipase C of Bacillus cereus.";
RL Gene 65:293-304(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VKM 164;
RX MEDLINE=93249510; PubMed=8387306;
RA Kuzmin N.P., Gavrilenko I.V., Krukov V.M., Karpov A.V.;
RT "Nucleotide sequence of phospholipase C and sphingomyelinase genes
RT from Bacillus cereus BKM-B164 (letter).";
RL Bioorg. Khim. 19:133-138(1993).
RN [3]
RP SEQUENCE OF 166-283 FROM N.A.
RC STRAIN=IAM 1208;
RX MEDLINE=88296483; PubMed=2841128;
RA Yamada A., Tsukagoshi N., Uda S., Sasaki T., Makino S., Nakamura S.,
RA Little C., Tomita M., Ikezawa H.;
RT "Nucleotide sequence and expression in Escherichia coli of the gene
RT coding for sphingomyelinase of Bacillus cereus.";
RL Eur. J. Biochem. 175:213-220(1988).
RN [4]
RP SEQUENCE OF 39-65.
RX MEDLINE=78043154; PubMed=72664;
RA Oltmanns A.-B., Little C., Sletten K., Wallin R., Johansen S.,
RA Flensburg R., Prydz H.;
RT "Some characteristics of phospholipase C from Bacillus cereus.";
RL Eur. J. Biochem. 79:459-468(1977).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=89159438; PubMed=2493587;
RA Hough E., Hansen L.K., Birkenes B., Jynge K., Hansen S., Hordvik A.,
RA Little C., Dodson E., Derewenda Z.;
RT "High-resolution (1.5 A) crystal structure of phospholipase C from
RT Bacillus cereus.";
RL Nature 338:357-360(1989).
CC -1- FUNCTION: REQUIRED, WITH SPHINGOMYELINASE TO EFFECT TARGET CELL
CC LYSIS (HEMOLYSIS).
CC -1- CATALYTIC ACTIVITY: A PHOSPHATIDYLCHOLINE + H(2)O = 1,2-
CC DIACYLGLYCEROL + CHOLINE PHOSPHATE.
CC -1- COFACTOR: THREE ZINC IONS.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: TO BETA-HEMOLYSIN OF STAPHYLOCOCCUS AUREUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X64141; CAA45502.1; -
DR EMBL: X12854; CAA31332.1; -
DR EMBL: X12711; CAA31213.1; -
DR EMBL: X64140; CAA45501.1; ALT-TERM.
DR PIR: S03200; S03200.
DR PIR: S01949; S01949.
DR PDB: 1AH7; 10-DEC-97.
DR InterPro: IPR001531; -
DR Pfam: PF00882; zn_dep_PLPC; 1.
DR PRINTS: PR00479; PRPHPLPASEC.
DR PROSITE: PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
KW Hydroxylase; Zinc; Signal; Zymogen; Hemolysis; 3D-structure.
FT SIGNAL 1 24
FT PROPEP 25 38
FT CHAIN 39 283 PHOSPHOLIPASE C.
FT METAL 39 39 ZINC 3.
FT METAL 52 52 ZINC 3.
FT METAL 93 93 ZINC 1.
FT METAL 107 107 ZINC 1.
FT METAL 156 156 ZINC 1.
FT METAL 160 160 ZINC 1 AND 3.
FT METAL 166 166 ZINC 2.
FT METAL 180 180 ZINC 2.
FT METAL 184 184 ZINC 2.
FT VARIANT 212 212 E -> D (IN STRAIN IAM1208).
FT VARIANT 226 226 S -> A (IN STRAIN IAM1208).
FT VARIANT 239 239 K -> R (IN STRAIN IAM1208).
FT VARIANT 282 282 D -> N (IN STRAIN IAM1208).
SQ SEQUENCE 283 AA; 32383 MW; ACS453EFP2E22B19 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 283;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DKMEAYITP 9
1111111111
DB 248 DKMEAYITP 256

PHLD_BACCE STANDARD: PRT: 283 AA.
ID PHLD_BACCE P33376;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE C PRECURSOR (EC 3.1.4.3) (PLC) (PHOSPHATIDYLCHOLINE
CHOLINEPHOSPHOHYDROLASE) (CEROLYSIN A).
GN CERA.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GP-4;
RX MEDLINE=89123149; PubMed=2536680;
RA Gilmore M.S., Cruz-Rodz A.L., Leimelster-Waechter M., Kreft J.,
Goebel W.;
RT "A Bacillus cereus cytolytic determinant, cerolysin AB, which
completes the phospholipase C and sphingomyelinase genes: nucleotide
sequence and genetic linkage.";
RL J. Bacteriol. 171:744-753(1989).
RN [2]
RP SEQUENCE OF 50-142 FROM N.A.
RA Gilmore M.S., Gilmore K.S., Goebel W.;
RT "A new strategy for ordered DNA sequencing based on a novel method for

the rapid purification of near-milligram quantities of a cloned
restriction fragment.";
RL Gene Anal. Tech. 2:108-114(1985).
CC -1- FUNCTION: REQUIRED, WITH SPHINGOMYELINASE TO EFFECT TARGET CELL
LYSIS (HEMOLYSIS).
CC -1- CATALYTIC ACTIVITY: A PHOSPHATIDYLCHOLINE + H(2)O = 1,2-
DIACYLGLYCEROL + CHOLINE PHOSPHATE.
CC -1- COFACTOR: THREE ZINC IONS.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: TO BETA-HEMOLYSIN OR STAPHYLOCOCCUS AUREUS.

DR EMBL: M24149; AAA91819.1; -
DR EMBL: M35411; AAA22524.1; -
DR PIR: A32042; A32042.
DR HSSP: P09596; 1AH7.
DR InterPro: IPR001531; -
DR Pfam: PF00882; zn_dep_PLPC; 1.
DR PRINTS: PR00479; PRPHPLPASEC.
DR PROSITE: PS00384; PROKAR_ZN_DEPEND_PLPC; 1.
KW Hydroxylase; Zinc; Signal; Zymogen; Hemolysis.
FT SIGNAL 1 24
FT PROPEP 25 38
FT CHAIN 39 283 BY SIMILARITY.
FT METAL 39 39 PHOSPHOLIPASE C.
FT METAL 52 52 ZINC 3.
FT METAL 93 93 ZINC 3.
FT METAL 107 107 ZINC 1.
FT METAL 156 156 ZINC 1.
FT METAL 160 160 ZINC 1.
FT METAL 166 166 ZINC 1 AND 3.
FT METAL 180 180 ZINC 2.
FT METAL 184 184 ZINC 2.
SQ SEQUENCE 283 AA; 32373 MW; 24D2E753402A6A44 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 283;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DKMEAYITP 9
1111111111
DB 248 DKMEAYITP 256

VPL_BPCHP STANDARD: PRT: 595 AA.
ID VPL_BPCHP P19192;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROTEIN VPI (ORF1).
OS Bacteriophage Chp1.
OC Viruses.
OX NCBI_TaxID=12367;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RX MEDLINE=90111716; PubMed=2607341;
RA Storey C.C., Lusher M., Richmond S.J.;
RT "Analysis of the complete nucleotide sequence of Chp1, a phage which
infects avian Chlamydia psittaci.";
RL J. Gen. Virol. 70:3381-3390(1989).
RN [2]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D00624; BAA00515.1; -
DR PIR: J00345; J00345.
KW Structural protein.
FT INIT_MET 0
FT SIMILAR 16 73 TO F PROTEIN OF PHAGE PH1-X174, AA
FT SEQUENCE 595 AA; 66850 MW; AE71D8024EA967ED CRC64;

Query Match 48.8%; Score 39; DB 1; Length 595;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEAYITPG 10
Db 93 DNWESFITGG 102

RESULT 11

SYT_MYCTU STANDARD; PRT; 1041 AA.
AC Q10765; O06181;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ISOEUCTL-TRNA SYNTHETASE (EC 6.1.1.5) (ISOEUCINE--TRNA LIGASE)
DE (ILERS).
GN ILES OR RV1536 OR MTCY48.29C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBL_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-ISOEUCINE + TRNA(ILE) = AMP +
CC PYROPHOSPHATE + L-ISOEUCTL-TRNA(ILE).
CC -1- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z74020; CAA98326.1; -
DR Tuberculin; Rv1536; -
DR InterPro: IPR001412; -
DR InterPro: IPR002300; -
DR InterPro: IPR002301; -
DR Pfam: PF00133; trna-synt_1; 1.

DR PRINTS: PR00984; TRNASYNTHILE.

DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Zinc.
FT SITE 53 63 "HIGH" REGION.
FT SITE 619 623 "RMSK" REGION.
FT BINDING 622 622 ATP (BY SIMILARITY).
FT SEQUENCE 1041 AA; 117339 MW; B5023822846E08C6 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 1041;
Best Local Similarity 62.5%; Pred. No. 66;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKWEAYIT 8
Db 134 DEMOAYVT 141

RESULT 12

RBBX_SCHPO STANDARD; PRT; 71 AA.
AC Q13877; O14458;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA-DIRECTED RNA POLYMERASES I, II, AND III 8.3 KDA POLYPEPTIDE
DE (EC 2.7.7.6) (ABCI0-BETA).
GN RPB10 OR SPAC1B3.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBL_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA MEDLINE=97177445; PubMed=9054344;
RA Shpakovskii G.V., Lebedenko E.N.;
RT "Three regions of Rpb10 mini-subunit of nuclear RNA polymerases are
RT strictly conserved in all eukaryotes (letter).";
RL Bioorg. Khim. 22:938-940(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA MEDLINE=98271918; PubMed=9608944;
RA Shpakovskii G.V., Proshkin S.A., Lebedenko E.N.;
RT "Exon-intron organization rpb10+ and rpb10+ genes of
RT Schizosaccharomyces pombe, coding for mini-subunits of nuclear RNA-
RT polymerase I-III.";
RL Mol. Biol. (Mosk) 32:285-290(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RT Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
CC RNA(N).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 15
CC DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS SHARED BY ALL 3 YEAST RNA
CC POLYMERASES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPB10 / EUKARYOTIC RPB10
CC RNA POLYMERASE SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF027818; AAC16895.1; -
 DR EMBL: D89596; BAA22805.1; -
 DR EMBL: U80219; AAC49842.1; -
 DR EMBL: Z98598; CAB1246.1; -
 DR InterPro: IPR000268; -
 DR Pfam: PF01194; RNA_POL_N; 1.
 DR PROSITE: PS01112; RNA_POL_N_8KD; 1.
 KW Transferrase: DNA-directed RNA polymerase; Transcription;
 KW Nuclear protein; Zinc; Metal-binding;
 FT METAL 7 7 ZINC (BY SIMILARITY).
 FT METAL 10 10 ZINC (BY SIMILARITY).
 FT METAL 44 44 ZINC (BY SIMILARITY).
 FT METAL 45 45 ZINC (BY SIMILARITY).
 SO SEQUENCE 71 AA; 8276 MW; 8B9EBE8FD59A2458 CRC64;

Query Match 47.5%; Score 38; DB 1; Length 71;
 Best Local Similarity 62.5%; Pred. No. 6.8;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DKWEAVIT 8
 |||:|:|
 Db 16 DKMDYLT 23

RESULT 13
 Y030_MYCTU STANDARD; PRT: 109 AA.
 AC P71600; 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HYPOTHETICAL 11.9 KDA PROTEIN RV0030.
 GN RV0030 OR MTCY10H4.30.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Ruter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z80233; CAB02415.1; -
 DR Tuberculin; RV0030; -
 KW Hypothetical protein.
 SO SEQUENCE 109 AA; 11858 MW; 11A342D48F17A0C7 CRC64;

Query Match 47.5%; Score 38; DB 1; Length 109;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DKWEAVIT 8
 |||:|:|
 Db 32 DKWEALVT 39

RESULT 14
 RS3A_SULSO STANDARD; PRT: 208 AA.
 ID RS3A_SULSO
 AC Q90XD4;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE 30S RIBOSOMAL PROTEIN S3AE.
 GN RPS3AE.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1617 / P2;
 RX MEDLINE=20165948; PubMed=10701121;
 RA Charlebois R.L., Singh R.K., Chan-Welther C.C.-Y., Allard G., Chow C.,
 RA Confalonieri F., Curtis R., Duguet M., Erauso G., Faguy D.,
 RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
 RA Kushwana N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
 RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 RA Ragan M.A., Senses C.W.;
 RT "Gene content and organization of a 281-kbp contig from the genome of
 RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
 RL Genome 43:116-136(2000).

CC -1 SIMILARITY: BELONGS TO THE S3AE FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Y18930; CAB57557.1; -
 DR InterPro: IPR001593; -
 DR Pfam: PF01015; Ribosomal_S3AE; 1.
 DR PROSITE: PS01191; RIBOSOMAL_S3AE; FALSE-NEG.
 KW Ribosomal protein.
 SO SEQUENCE 208 AA; 23542 MW; 3837ABE7E53676B6 CRC64;

Query Match 47.5%; Score 38; DB 1; Length 208;
 Best Local Similarity 63.6%; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 KWEAVITPGAF 12
 ||:|||||
 Db 16 KMTSVITPKAF 26

RESULT 15
 SSRA_ARATH STANDARD; PRT: 257 AA.
 ID SSRA_ARATH
 AC P43434; Q9SKP7;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TRANSLACON-ASSOCIATED PROTEIN, ALPHA SUBUNIT PRECURSOR (TRAP-ALPHA)
 DE (SIGNAL SEQUENCE RECEPTOR ALPHA SUBUNIT) (SSR-ALPHA).
 GN ATG21160 OR F26H11.8.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID:3702;

Search completed: November 19, 2001, 08:25:35
 Job time: 723 sec

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RX MEDLINE-20083487; PubMed-10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bento M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Niernm W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE OF 5'-257 FROM N.A.
 RX MEDLINE-94326944; PubMed-8050590;
 RA Hartmann E., Prehn S.;
 RT "The N-terminal region of the alpha-subunit of the TRAP complex has a
 conserved cluster of negative charges.";
 RL FEBS Lett. 349:324-326(1994).
 CC -1- FUNCTION: TRAP PROTEINS ARE PART OF A COMPLEX WHOSE FUNCTION IS TO
 BIND CA(2+) TO THE ER MEMBRANE AND THEREBY REGULATE THE RETENTION
 OF ER RESIDENT PROTEINS. MAY BE INVOLVED IN THE RECYCLING OF THE
 TRANSLLOCATION APPARATUS AFTER COMPLETION OF THE TRANSLLOCATION
 PROCESS OR MAY FUNCTION AS A MEMBRANE-BOUND CHAPERONE FACILITATING
 FOLDING OF TRANSLLOCATED PROTEINS.
 CC -1- SUBUNIT: HETEROTETRAMER OF TRAP-ALPHA, TRAP-BETA, TRAP-DELTA AND
 TRAP-GAMMA.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ENDOPLASMIC
 RETICULUM.
 CC -1- DOMAIN: SHOWS A REMARKABLE CHARGE DISTRIBUTION WITH THE N-TERMINUS
 BEING HIGHLY NEGATIVELY CHARGED, AND THE CYTOPLASMIC C-TERMINUS
 POSITIVELY CHARGED.
 CC -1- PTM: PHOSPHORYLATED IN ITS CYTOPLASMIC TAIL (BY SIMILARITY).
 CC -1- MISCELLANEOUS: SEEMS TO BIND CALCIUM.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AC006264; AAD29800.1; -;
 DR EMBL; L32016; AAA21820.1; -;
 KW Glycoprotein; Signal; Phosphorylation; Endoplasmic reticulum;
 KW Transmembrane; Calcium-binding;
 FT SIGNAL 1 23
 FT CHAIN 24 257
 FT POTENTIAL.
 FT TRANSLOCON--ASSOCIATED PROTEIN, ALPHA
 FT SUBUNIT.
 FT LUMENAL (POTENTIAL).
 FT DOMAIN 24 189
 FT TRANSMEM 190 208
 FT POTENTIAL.
 FT DOMAIN 209 257
 FT CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 56 56
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 118 118
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 257 AA; 28036 MW; C43CEDCAC80D2659 CRC64;

Query Match 47.5%; Score 38; DB 1; Length 257;
 Best Local Similarity 66.7%; Pred. No. 24;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 YITPGAFDV 14
 1:|||||
 DB 146 YITPGAFDL 154

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:01 ; Search time 81.19 Seconds
(without alignments)
22.814 Million cell updates/sec

Title: US-09-610-118-63

Perfect score: 80

Sequence: 1 DKWEAYITPGAPDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL_16:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_unclassified:*

13: sp_vertebrate:*

14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	58.8	325	5 09M4N0	09M4N0 drosophila
2	47	58.8	325	5 024157	024157 drosophila
3	45	56.2	765	2 P74599	P74599 synechocyst
4	44	55.0	175	9 080091	080091 bacterioph
5	44	55.0	354	5 09VCM5	09VCM5 drosophila
6	44	55.0	669	2 070012	070012 streptomyc
7	44	55.0	669	2 09KX04	09KX04 streptomyc
8	44	55.0	682	4 043395	043395 homo sapien
9	44	55.0	683	4 043446	043446 homo sapien
10	41	51.2	333	1 09HLB9	09HLB9 thermoplasm
11	41	51.2	312	5 09G116	09G116 plasmodium
12	41	51.2	422	10 09RSY1	09RSY1 deinococcus
13	41	51.2	422	10 039811	039811 glycine max
14	41	51.2	446	2 09RI08	09RI08 streptomyc
15	41	51.2	684	2 09KJ70	09KJ70 vibrio chol
16	41	51.2	741	10 09FJD2	09FJD2 arabisdopsi
17	41	51.2	948	6 09TMD7	09TMD7 trichosurus
18	41	51.2	1846	5 09VMN8	09VMN8 drosophila
19	40	50.0	249	2 09EX43	09EX43 streptomyc

20	40	50.0	267	10 09FYR8	09FYR8 oryza sativ
21	40	50.0	365	10 09FT98	09FT98 arabisdopsi
22	40	50.0	417	10 09FXT4	09FXT4 oryza sativ
23	40	50.0	539	2 09RDS9	09RDS9 streptomyc
24	40	50.0	721	1 050117	050117 pyrococcus
25	40	50.0	2122	2 09XBC2	09XBC2 amycolatops
26	40	50.0	3158	2 052819	052819 amycolatops
27	40	48.8	148	1 059589	059589 pyrococcus
28	39	48.8	156	2 09S6G4	09S6G4 streptococc
29	39	48.8	155	2 09S6G3	09S6G3 streptococc
30	39	48.8	179	2 09ZMT6	09ZMT6 helicobacte
31	39	48.8	217	2 09XSC4	09XSC4 streptococc
32	39	48.8	264	2 053751	053751 mycobacteri
33	39	48.8	268	2 09XSC6	09XSC6 streptococc
34	39	48.8	283	2 052864	052864 bacillus th
35	39	48.8	434	10 09LY12	09LY12 arabisdopsi
36	39	48.8	495	2 0910S2	0910S2 pseudomonas
37	39	48.8	684	10 09SBC6	09SBC6 arabisdopsi
38	39	48.8	766	10 082485	082485 proteus mir
39	39	48.8	843	2 P72211	P72211 proteus mir
40	41	48.8	921	10 09FMC5	09FMC5 oryza sativ
41	39	48.8	998	10 09SVG2	09SVG2 arabisdopsi
42	39	48.8	1734	5 09U0L7	09U0L7 plasmodium
43	39	48.8	3306	10 09FT44	09FT44 arabisdopsi
44	38	47.5	98	10 09M5V3	09M5V3 glycine max
45	38	47.5	204	10 004944	004944 hordeum vul

ALIGNMENTS

RESULT 1

ID 09M4N0 PRELIMINARY: PRT: 325 AA.

AC 09M4N0: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE BRN PROTEIN.
GN BRN OR EG:EG0007.6 OR CG4934.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arif J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Itoigawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenson J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,


```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98067870; PubMed-9404084;
RA Kaneko J., Kimura T., Kawakami Y., Tomita T., Kamio Y.;
RT "Pancon-valentine leukocidin genes in a phage-like particle isolated
RL from milcomycin C-treated Staphylococcus aureus V8 (ATCC 49775).";
DR EMBL; AB009866; BAA31926.1;
DR HSSP; P16088; 1DUF;
DR InterPro: IPR001428;
DR Pfam: PF00692; d0rpfase; 1.
DR ProDom: PD000946;
SQ SEQUENCE 175 AA; 19081 MW; B93B547104EC9BA0 CRC64;

Query Match 55.0%; Score 44; DB 9; Length 175;
Best Local Similarity 77.8%; Pred. No. 7.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 YTPGADV 14
Db 108 YTPGADV 116

RESULT 5
Q9VCT5 PRELIMINARY; PRT; 354 AA.
AC Q9VCT5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2001 (TrEMBLrel. 16, Last annotation update)
DE C64704 PROTEIN.
GN C64704.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID:7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkov D., Botchan M.R., Bouck J., Brothstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de la Piedad B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodali C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McInerney D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```

```

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: TO EF-HAND FAMILY.
DR EMBL; AE003740; AAF56070.1;
DR FlyBase: FBgn0039029; C64704.
DR InterPro: IPR002048;
DR Pfam: PF00036; ehand; 2.
DR PROSITE: PS00018; EF_HAND; 2.
DR SMART: SM00054; EFh; 1.
DR K Calcium-binding.
SQ SEQUENCE 354 AA; 41790 MW; 9B00F103E6EC6B1 CRC64;

Query Match 55.0%; Score 44; DB 5; Length 354;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAVTPGAF 12
Db 31 KWEAVTPGAF 41

RESULT 6
O70012 PRELIMINARY; PRT; 669 AA.
ID O70012;
AC O70012;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PEPL.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID:1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Schneider D., Bruton C.J., Chater K.F.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE-96134943; PubMed-8596463;
RA Bruton C.J., Plaskitt K.A., Chater K.F.;
RT "Tissue-specific glycogen branching isoenzymes in a multicellular
RL prokaryote, Streptomyces coelicolor A3(2).";
DR EMBL; AJ001206; CAA04606.1;
DR InterPro: IPR000461;
DR InterPro: IPR000531;
DR InterPro: IPR001589;
DR Pfam: PF00128; alpha-amylase; 1.
DR PROSITE: PS00019; ACTININ.1; UNKNOWN.1.
DR PROSITE: PS00430; TONB-DEPENDENT_REC.1; UNKNOWN.1.
SQ SEQUENCE 669 AA; 74113 MW; 153E0134CD16268 CRC64;

Query Match 55.0%; Score 44; DB 2; Length 669;
Best Local Similarity 63.6%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEAVTPGA 11
Db 77 DRWGATVPGA 87

RESULT 7
Q9KT04

```

```

ID 09KY04 PRELIMINARY: PRT: 669 AA.
AC 09KY04;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PUTATIVE ALPHA-AMYLASE.
GN PEP1B.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Streptomyces.
OC Actinomycetales; Streptomyces; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RC Seeger K.J., Harris D.;
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RC Thomson N.R., Parkhill J., Barrell B.G., Raftery M.A.;
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RC MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleser H.M., Denaplatte D., Elchner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmid and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL356932; CAB92881.1; -.
DR InterPro: IPR000531; -.
DR PROSITE: PS00019; ACTININ.1; UNKNOWN_1.
DR PROSITE: PS00430; TONB DEPENDENT REC.1; UNKNOWN_1.
SQ SEQUENCE 669 AA; 7413 MW; 5463E0134CD6518 CRC64;

Query Match 55.0%; Score 44; DB 2; Length 669;
Best Local Similarity 63.6%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DKWEYITPGA 11
DB 77 DRWGAITPGA 87

RESULT 8
ID 043395 PRELIMINARY: PRT: 682 AA.
AC 043395;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE U4/U6-ASSOCIATED RNA SPLICING FACTOR.
GN PRP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97472464; PubMed=9328476;
RA Wang A., Forman-Kay J., Luo Y., Luo M., Chow Y.H., Plumb J.,
RA Friesen J.D., Tsui L.C., Heng H.H., Woolford J.L. Jr., Hu J.;
RT "Identification and characterization of human genes encoding Hprp3
RT and Hprp4, interacting components of the spliceosome."
RL Hum. Mol. Genet. 6:2117-2126(1997).
DR EMBL: AF001947; AAC09069.1; -.
DR InterPro: IPR002483; -.
DR Pfam: PF01480; PMI. 1.
DR SMART: SM00311; PMI. 1.
SQ SEQUENCE 682 AA; 77403 MW; 544660BF72A39BA3 CRC64;

```

```

Query Match 55.0%; Score 44; DB 4; Length 682;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYTPGAFDV 14
DB 390 WDSYIIPNGFDL 401

RESULT 9
ID 043446 PRELIMINARY: PRT: 683 AA.
AC 043446;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE U4/U6 SMALL NUCLEAR RIBONUCLEOPROTEIN HPRP3.
GN HPRP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98067393; PubMed=9404889;
RA Horowitz D.S., Kobayashi R., Krainer A.R.;
RT "A new cyclophilin and the human homologues of yeast Prp3 and Prp4
RT form a complex associated with U4/U6 snRNPs."
RL RNA 3:1374-1387(1997).
DR EMBL: AF016370; AAC51926.1; -.
DR InterPro: IPR002483; -.
DR Pfam: PF01480; PMI. 1.
DR SMART: SM00311; PMI. 1.
DR Nucleotide; Ribonucleoprotein.
SQ SEQUENCE 683 AA; 77528 MW; 4AA6AA4C99110284 CRC64;

Query Match 55.0%; Score 44; DB 4; Length 683;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 WEAYTPGAFDV 14
DB 390 WDSYIIPNGFDL 401

RESULT 10
ID 09HLB9 PRELIMINARY: PRT: 333 AA.
AC 09HLB9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE UROPOPHYRINOGEN DECARBOXYLASE RELATED PROTEIN.
GN TA0310.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
OC Thermoplasmata.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM 1728;
RC MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frisman D., Stocker S., Lups A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
DR EMBL: AL445063; GAC11455.1; -.
DR InterPro: IPR000257; -.
DR InterPro: IPR000652; -.

```

DR Pfam: PF01208; URO-D; 1.
DR ProDom: PD001005; -. 1.
DR PROSITE: PS00906; UROD_1; UNKNOWN_1.
SO SEQUENCE 333 AA; 37785 MW; 3DBE510768C0F5B CRC64;

Query Match 53.8%; Score 43; DB 1; Length 333;
Best Local Similarity 38.5%; Pred. No. 24;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
| | | | | : : :
Db 199 DSMAGTISPEYE 211

RESULT 11

096116 PRELIMINARY; PRT; 312 AA.
AC 096116;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RIFIN.
GN PERO055C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551.
RA Gardner M.J., Tetzelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shalimov S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Perlea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
falciparum".
RL Science 282:1126-1132(1998).
DR EMBL: AE001368; AAC71801.1; -.
DR InterPro: IPR002858; -.
DR Pfam: PF02099; R1fln_STEVOR; 1.
SO SEQUENCE 312 AA; 35101 MW; BAF5EDB2BE1957D2 CRC64;

Query Match 51.2%; Score 41; DB 5; Length 312;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 KWEAYITPGAFD 13
| | | | | :
Db 177 KWSAMVTPPTVD 188

RESULT 12

09RSY1 PRELIMINARY; PRT; 380 AA.
AC 09RSY1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2001 (TREMBLrel. 16, Last annotation update)
DE PHOH-RELATED PROTEIN.
GN DR1988.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1".
RL Science 286:1571-1577(1999).
DR EMBL: AE002037; AAF11540.1; -.
DR TIGR; DR1988; -.
DR InterPro: IPR000958; -.
DR SMART: SM00322; KH; 1.
SO SEQUENCE 380 AA; 41377 MW; C949F78120CBF87B CRC64;

Query Match 51.2%; Score 41; DB 2; Length 380;
Best Local Similarity 42.9%; Pred. No. 61;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFDV 14
| | | | | : :
Db 243 KEFAYLTSGVIEI 256

RESULT 13

039811 PRELIMINARY; PRT; 422 AA.
AC 039811;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE ALPHA GALACTOSIDASE.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WILLIAMS;
RA Davis M.O., Walker J.C., Smith D.;
RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL: U12926; AAA73963.1; -.
DR Mendel: 8332; Glyma.1501; 8332.
DR InterPro: IPR00111; -.
DR InterPro: IPR002241; -.
DR InterPro: IPR002287; -.
DR Pfam: PF02065; Melibiose; 2.
DR PRINTS: PR00740; GLHYDRASE27.
DR ProDom: PD002572; -. 1.
DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
SO SEQUENCE 422 AA; 46395 MW; 9281B25EAD6940A0 CRC64;

Query Match 51.2%; Score 41; DB 10; Length 422;
Best Local Similarity 46.2%; Pred. No. 69;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
| | | | | : : :
Db 260 DKWASYAGRGGMN 272

RESULT 14

09RIU8 PRELIMINARY; PRT; 446 AA.
AC 09RIU8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHEICAL 47.1 KDA PROTEIN.
GN SCM1.13C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2).
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2).
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2).
RX MEDLINE=97000351; Pubmed=8843436;
RA Keeneyach M., Kleiser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.",
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL133278; CAB61924.1; -.
KW Hypothetical protein.
SQ SEQUENCE 446 AA; 47146 MW; 0F5A08345C500576 CRC64;

Query Match	51.2%	Score 41:	DB 2:	Length 446:
Best Local Similarity	66.7%	Pred. No. 73:		
Matches	6:	Conservative	2:	Mismatches
			1:	Indels
				Gaps
Qy	3 WEATITPCA 11			
	:			
Db	121 WEAHVLPQA 129			

Best Local Similarity	53.3%	Pred. No. 1.2e+02;	
Matches	8;	Conservative	2;
		Mismatches	1;
		Indels	4;
		Gaps	1.

Search completed: November 19, 2001, 08:25:02
Job time: 740 sec

Query Match	51.28; Score 41; DB 2; Length 684;
-------------	------------------------------------

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:12:34 ; Search time 77.92 Seconds
(without alignments)
10.892 Million cell updates/sec

Title: US-09-610-118-63

Perfect score: 80

Sequence: 1 DKWEAYITPGAFDV 14

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: /SIDSB8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSB8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSB8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSB8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSB8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSB8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSB8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSB8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSB8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSB8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSB8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSB8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSB8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSB8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSB8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSB8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSB8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSB8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSB8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSB8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSB8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDSB8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	14	22	AAB61295 Anti-TANGO 268 scf
2	44	55.0	169	18	AAW28034 Amino acid sequenc
3	43	53.8	207	21	AAG34190 zea mays protein f
4	43	53.8	217	21	AAG34189 zea mays protein f
5	43	53.8	223	21	AAG34188 zea mays protein f
6	41	51.2	312	21	AAB18153 Plasmodium falcipa
7	41	51.2	363	18	AAW10320 soybean alpha-D-ga
8	41	51.2	378	16	AAW70205 Alpha-galactosidas
9	41	51.2	378	17	AAW00621 Coffee bean alpha
10	41	51.2	406	16	AAW26604 Sena alpha-galact
11	41	51.2	411	16	AAW70206 Alpha-galactosidas

12	41	51.2	420	16	AAW71326 Coffee bean alpha-
13	40	50.0	116	15	AAW52061 Heavy chain variab
14	40	50.0	117	15	AAW52065 Heavy chain variab
15	40	50.0	117	15	AAW52067 Heavy chain variab
16	40	50.0	221	21	AAB18667 A human regulator
17	40	50.0	365	21	AAW20888 Arabidopsis thalia
18	40	50.0	367	21	AAW20887 Arabidopsis thalia
19	40	50.0	366	21	AAW20886 Arabidopsis thalia
20	39	48.8	141	21	AAW71040 Streptococcus pyog
21	39	48.8	155	21	AAW71044 Streptococcus pyog
22	39	48.8	159	21	AAW71041 Streptococcus pyog
23	39	48.8	167	21	AAW71046 Streptococcus pyog
24	39	48.8	179	18	AAW55630 H. pylori ORF 02ae
25	39	48.8	184	18	AAW55230 H. pylori ORF 02ce
26	39	48.8	217	21	AAW71039 Streptococcus pyog
27	39	48.8	217	21	AAW71042 Streptococcus pyog
28	39	48.8	271	21	AAW71045 Streptococcus pyog
29	39	48.8	766	21	AAW96161 Arabidopsis thalia
30	39	48.8	1045	19	AAW57448 M. tuberculosis is
31	38	47.5	114	21	AAW51263 Monoclonal antibod
32	38	47.5	117	11	AAW04381 Variable heavy cha
33	38	47.5	117	14	AAW04380 Sequence of VH-41P
34	38	47.5	117	20	AAW57175 Amino acid sequenc
35	38	47.5	117	20	AAW50683 Human VH alphasg
36	38	47.5	117	21	AAW90714 Vha1phatag VH regl
37	38	47.5	117	21	AAW57046 Vha1phatag amino a
38	38	47.5	133	10	AAW90844 Predicted sequence
39	38	47.5	133	11	AAW04384 Sequence of murine
40	38	47.5	133	14	AAW38313 Amino acid sequenc
41	38	47.5	133	20	AAW57177 Human CC83 protein
42	38	47.5	133	20	AAW50686 CC83 VH region pro
43	38	47.5	133	21	AAW90716 CC83 VH region CNB
44	38	47.5	133	21	AAW90723 Amino acid sequenc
45	38	47.5	133	21	AAW57048

ALIGNMENTS

RESULT 1	
ID	AAW61295 standard; Peptide: 14 AA.
XX	
AC	AAW61295:
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Anti-TANGO 268 scFv CDR, SEQ ID NO: 63.
XX	
KW	Human; antibody; scFv; CDR; complementarity determining region;
KW	TANGO 268; cardiant; cerebroprotective; cytostatic; anticoagulant;
KW	thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;
KW	platelet membrane glycoprotein receptor; bleeding disorder;
KW	blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;
KW	ischemia; cardiovascular disease; immunological disease; liver disorder;
KW	Cancer.
OS	
XX	
OS	Homo sapiens.
XX	
PN	WO200100810-A1.
XX	
PD	
XX	
XX	04-JAN-2001.
XX	
XX	30-JUN-2000; 2000WO-US18152.
XX	
PF	
XX	
PR	30-JUN-1999; 99US-0345468.
XX	
PR	06-DEC-1999; 99US-0454824.
XX	
PR	14-FEB-2000; 2000US-0503387.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Bufield SJ, Villetal J, Jandrot-Perrus M, Valinchenker W, Gill DS;
PI	Qian MD, Kingsbury G;

XX WPI; 2001-080877/09.
DR
XX
PT New genes encoding human platelet-expressed collagen receptor,
PT glycoprotein VI, and its modulators, useful for preventing, treating
PT and diagnosing hemorrhagic disorders, thrombotic diseases and
PT immunological disorders -
XX
XX Claim 31; Page 102; 227pp; English.
XX
XX The present sequence is given in a specification relating to an isolated
CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
CC and polypeptides and their modulators, e.g. antisense nucleic acids,
CC ribozymes and antibodies, are useful for preventing, treating and
CC diagnosing disorders associated with aberrant expression or activity of
CC GPVI. These disorders include bleeding disorders
CC (e.g. thrombocytopaenia), blood vessel injury, thrombotic disorders
CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC infarction), immunological diseases (e.g. platelet disorder) and
CC embryonic liver disorders. Preferably they are used to prevent acute
CC cardiac ischaemia following angioplasty and metastatic cancers,
CC especially of the colon and liver.
XX
XX Sequence 14 AA:
SQ

Query Match 100.0%; Score 80; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. NO. 1.8e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFDV 14
Db 1 dkweayitpgadv 14
|||||

RESULT 2
AAW28034 standard; Protein; 169 AA.
XX
XX AAW28034;
XX
DT 27-AUG-1998 (first entry)
XX
XX Amino acid sequence of the deoxyuridine 5' triphosphatase homologue.
DE
XX
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome.
XX
XX Staphylococcus aureus.
OS
XX
XX WO9730070-A1.
PN
XX
XX 21-AUG-1997.
PD
XX
XX 19-FEB-1997; 97WO-US02318.
PF
XX
XX 20-FEB-1996; 96US-0011888.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA
XX
XX Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX
XX WPI: 1997-424969/39.
DR
XX
XX M-PSDB: AAT83988.
PT
XX
XX Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
PT to isolate antimicrobial compounds, and in vaccines against S.

PT aureus infection
XX
XX
PS Claim 6; Page 421; 989pp; English.
XX
XX The present sequence represents a Staphylococcus aureus protein,
CC which is believed to be a deoxyuridine 5' triphosphatase homologue.
CC The DNA sequence was isolated from a library of clones of S. aureus
CC WCUH 29 in Escherichia coli. The DNA sequence can be used in the
CC construction of ribozymes and antisense sequence to control the
CC expression of Staphylococcal genes. The DNA sequence is also useful as
CC a source of regulatory elements for the control of bacterial gene
CC expression. The present protein may be used to produce vaccines to
CC enable a host to produce specific antibodies with antibacterial action.
CC These vaccines and antibodies would protect a host against invasion by
CC S. aureus, and conditions relating to Staphylococcal infection, e.g.
CC Staphylococcal food poisoning, scaled skin syndrome, and toxic shock
CC syndrome.
XX
XX Sequence 169 AA:
SQ

Query Match 55.0%; Score 44; DB 18; Length 169;
Best Local Similarity 77.8%; Pred. NO. 4.5;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 YITPGAFDV 14
Db 102 yitpgadvf1 110
|||||

RESULT 3
AAG34190 standard; Protein; 207 AA.
ID AAG34190
XX
XX AAG34190;
XX
DT 18-OCT-2000 (first entry)
XX
XX Zea mays protein fragment SEQ ID NO: 41561.
DE
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
XX Zea mays subsp. mays.
OS
XX
XX EP1033405-A2.
FN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128214.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136332.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

```
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match          53.8%; Score 43; DB 21; Length 207;
Best Local Similarity 42.9%; Pred. No. 8.3;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Caps 0;

OY 1 DKWEAYITPGAFDV 14
   :||| :|| :|
Db 83 ewealllpeqfnl 96

RESULT 4
AACG34189
ID AACG34189 standard; Protein; 217 AA.
XX AACG34189;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Zea mays protein fragment SEQ ID NO: 41560.
DE
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
XX Zea mays subsp. mays.
OS
XX
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
```

```
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141267.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 53.8%; Score 43; DB 21; Length 217;
Best Local Similarity 42.9%; Pred. No. 8.7;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 1 DKWEATTPGAFDV 14
Db 93 eeewalltpedfnl 106
RESULT 5
AAG34188
ID AAG34188 standard; Protein: 223 AA.
XX
AC AAG34188;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 41559.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
FN Ep1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0153363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157755.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.


```
XX Alpha-galactosidase of green coffee bean.
DE
XX
KM Alpha-galactosidase; coffee bean; CBG; alpha-1,3-linked galactose;
KM B antigen; blood group; Sfg.
XX
XX Coffee sp.
OS
FH Key Location/Qualifiers
FT Peptide 1..15
FT /label= Sig-peptide
XX
XX WO9507088-A.
PN
XX
XX 16-MAR-1995.
PD
XX
XX 26-AUG-1994; 94WO-US09662.
PF
XX 08-SEP-1993; 93US-0118470.
PR
XX
XX (NYBL-) NEW YORK BLOOD CENT INC.
PA
XX Goldstein J, Zhu A;
PI
XX WPI: 1995-123231/16.
DR
XX N-PSDB; AA083523.
XX
XX New recombinant coffee bean alpha-galactosidase - used for
PT cleaving alpha1,3-linked galactose residues on the surface of
PT cells for prodn. of blood prods.
XX
XX Disclosure: Page 38-39; 67pp; English.
PS
XX
XX DNA encoding coffee bean alpha-galactosidase (CBG) was obtained by
CC isolating mRNA from coffee beans, prep. cDNA and subjecting this
CC to PCR amplification using primers based on the amino acid
CC sequence of CBG. A full-length cDNA clone is given in AA083523, which
CC encoded a 42 kDa protein (AA070205). Recombinant CBG was produced in
CC Sf9 cells.
XX
XX Sequence 378 AA:
SQ

Query Match 51.2%; Score 41; DB 16; Length 378;
Best Local Similarity 46.2%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 DKWEAYITPGAFD 13
DB 216 dkwasayagpggwn 228

RESULT 9
AAW00621
ID AAW00621 standard; Protein: 378 AA.
XX
XX AAW00621;
AC
XX 15-NOV-1996 (first entry)
DT
XX
XX Coffee bean alpha-galactosidase.
DE
XX
XX Alpha-galactosidase; coffee; blood group; B antigen;
KM Pichia pastoris.
XX
XX Coffee sp.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..15
FT /label= Sig-peptide
FT Modified-site 151..153
FT /label= N-linked_glycosylation_site
XX
```

```
PN WO9523869-A1.
XX
XX 08-AUG-1996.
PD
XX
XX 30-JAN-1996; 96WO-US01212.
PF
XX
XX 30-JAN-1995; 95US-0380194.
PR
XX
XX (NYBL-) NEW YORK BLOOD CENT INC.
PA
XX Goldstein J, Zhu A;
PI
XX
XX WPI: 1996-371420/37.
DR
XX N-PSDB; AAT35799.
XX
XX New recombinant coffee bean alpha-galactosidase - used partic. for
PT removing B antigens from the surface of cells in blood prods.
PT
XX
XX Example 1; Fig 1A-B; 56pp; English.
PS
XX
XX The amino acid sequence (AAW00621) of coffee bean full-length alpha-
CC galactosidase was deduced from a cDNA clone (AAT35799) obtd. from
CC dried green coffee beans. Large-scale prodn. of recombinant enzyme
CC was achieved using Pichia pastoris as host. The enzyme can be used
CC to cleave galactose sugar residues, partic. alpha-1,3-linked
CC galactose residues. It is useful for the removal of type B antigens
CC from the surface of cells in blood prods., thereby converting type
CC B blood prods. to type O, and type AB to type A without affecting
CC the surface or function of the cells.
XX
XX Sequence 378 AA:
SQ

Query Match 51.2%; Score 41; DB 17; Length 378;
Best Local Similarity 46.2%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 DKWEAYITPGAFD 13
DB 216 dkwasayagpggwn 228

RESULT 10
AAW26604
ID AAW26604 standard; Protein: 406 AA.
XX
XX AAW26604;
AC
XX 27-JAN-1998 (first entry)
DT
XX
XX Senna alpha-galactosidase.
DE
XX
XX Alpha-galactosidase; mannanose; galactose; galactomannan; guar gum;
KM locust bean gum; transgenic plant.
XX
XX Senna.
OS
XX
XX WO9720937-A2.
PN
XX 12-JUN-1997.
PD
XX
XX 02-DEC-1996; 96WO-EP05581.
PF
XX
XX 04-DEC-1995; 95GB-0024752.
PR
XX
XX (DANI-) DANISCO AS.
PA
XX Brunstedt J, Jorsboe M, Petersen SG;
PI
XX WPI: 1997-319783/29.
DR
XX N-PSDB; AAT90432.
XX
XX In vivo modification process that affects the mannanose-to-galactose
```

PT ratio of an organism - useful to convert guar gum into locust bean gum

PS Claim 14; Fig 4; 73pp; English.

XX This polypeptide sequence comprises senna alpha-galactosidase.

CC The sequence was deduced from PCR clones (see AAR70431) obtained from

CC endosperm cDNA, and deposited in E. coli K12 as NCIMB 40831. The

CC alpha-galactosidase can be used in a claimed *in vivo* modification

CC process to increase the mannose-to-galactose ratio of either an

CC organism, preferably a transgenic guar plant, or of a mannose/

CC galactose-containing compound such as guar gum galactomannan. In

CC *vivo* modified guar gum with improved functional properties, such as

CC a mannose-to-galactose ratio similar to that of locust bean gum,

CC can be obtained for use in the food industry. Improved

CC galactomannans have also been produced using a cloned guar

CC phosphomannose isomerase (see AAR70603).

XX Sequence 406 AA;

SO

Query Match 51.2%; Score 41; DB 18; Length 406;

Best Local Similarity 46.2%; Pred. No. 38;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKMEAYITPGAFD 13

Db 244 dkwasayagpgwn 256

|||||

RESULT 11

AAR70206

ID AAR70206 standard; Protein; 411 AA.

XX AAR70206;

XX 21-SEP-1995 (first entry)

DE Alpha-galactosidase of guar.

XX

KM Alpha-galactosidase; coffee bean; CBG: alpha-1,3-linked galactose;

KM B antigen; blood group; guar.

XX

OS Cymopsis tetragonoloba.

XX

PN W09507088-A.

XX

PD 16-MAR-1995.

XX

PF 26-AUG-1994; 94WO-US09662.

XX

PR 08-SEP-1993; 93US-0118470.

XX

PA (NYBL-) NEW YORK BLOOD CENT INC.

XX

PI Goldstein J, Zhu A;

XX

DR WPI: 1995-123231/16.

XX

PT New recombinant coffee bean alpha-galactosidase - used for

PT cleaving alpha1,3-linked galactose residues on the surface of

PT cells for prodn. of blood prods.

XX

PS Disclosure: Page 42-43; 67pp; English.

XX

CC DNA encoding coffee bean alpha-galactosidase (CBG) was obtained by

CC isolating mRNA from coffee beans, prep. cDNA and subjecting this

CC to PCR amplification using primers based on the amino acid

CC sequence of CBG. A full-length cDNA clone is given in AA083523, which

CC encoded a 42 kDa protein (AAR70205). Recombinant CBG was produced in

CC Sf9 cells. CBG showed homology to alpha-galactosidases from guar,

CC human placenta, *S. cerevisiae* and *A. niger* (AAR70206-09).

XX

SO Sequence 411 AA;

Query Match 51.2%; Score 41; DB 16; Length 411;

Best Local Similarity 46.2%; Pred. No. 39;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKMEAYITPGAFD 13

Db 248 dkwasayagpgwn 260

|||||

RESULT 12

AAR71326

ID AAR71326 standard; Protein; 420 AA.

XX AAR71326;

XX 18-NOV-1995 (first entry)

DE Coffee bean alpha-galactosidase.

XX

KM Alpha-galactosidase; enzyme.

XX

OS Coffee.

XX

FH Key Location/Qualifiers

FT Peptide 1..38

FT Protein 39..420

FT Protein 58..420

FT Protein 58..420

FT Cleavage-site 38..39

FT Cleavage-site 193

FT Modified-site 193

FT /label= glycosylation site

XX

PN W09506478-A.

XX

PD 09-MAR-1995.

XX

PF 30-AUG-1994; 94WO-US09739.

XX

PR 30-AUG-1993; 93US-0113890.

XX

PA (HAMA-) HAMATI BIOTECHNOLOGY GROUP INC.

XX

PI Clements DE, Ivy JM;

XX

DR WPI: 1995-115260/15.

XX

DR N-PSDB; AA085644.

XX

PT DNA encoding coffee bean alpha-galactosidase - also the

PT recombinant enzyme, useful for converting human type B

PT erythrocytes to type O

XX

PS Claim 7; Figure 1A-1B; 40pp; English.

XX

CC Mature coffee bean alpha-galactosidase is produced as a preproenzyme

CC (see AAR71326 FT). Analogies with other preproenzymes suggest a signal

CC protease cleavage site as indicated in FT. The mature protein

CC contains 363 AAs, as verified by N-terminal sequencing. RNA from

CC stages 5 and 6 coffee beans was used to construct a cDNA library

CC which was screened. Clone p27alpagc1-1 was sequenced (see AA085644).

XX

SO Sequence 420 AA;

Query Match 51.2%; Score 41; DB 16; Length 420;

Best Local Similarity 46.2%; Pred. No. 40;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFD 13
111:111:
DB 258 dkwaysy9p9gwn 270

RESULT 13

ARS2061
ID ARS2061 standard; Protein; 116 AA.

AC AAR52061;

DT 11-OCT-1996 (first entry)

DE Heavy chain variable region of murine anti-N901 antibody.

KW antibody; humanised; murine; human; heavy chain; light; variable;
framework region; complementarity determining region; reshaping;
modelling; surface residue; modify.

OS Mus sp.

Key Location/Qualifiers

FT 1..30 /label= framework_region_1

FT /note= "FR 1"

FT Region

FT 31..35 /label= complementarity_determining_region_1

FT /note= "CDR 1"

FT Region

FT 36..49 /note= "FR 2"

FT Region

FT 50..59 /note= "CDR 2"

FT Region

FT 60..98 /note= "FR 3"

FT Region

FT 99..107 /note= "CDR 3"

FT Region

FT 13-APR-1994.

PD 07-SEP-1993; 93EP-0307051.

PF 09-SEP-1992; 92US-0942245.

PR (PEDE/) PEDERSEN J T.

PA (IMMU-) IMMUNOGEN INC.

XX Guld BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;

XX WPI: 1994-120230/15.

DR Method of resurfacing of rodent antibodies to produce humanised

PT antibody forms - for producing non-human antibodies with improved

PT therapeutic efficiency by presenting human surface on V-region

PT Example 1; Fig 4B; 230pp; English.

PS Modification of a rodent antibody or fragment by resurfacing in order

XX to produce a humanised rodent antibody can be determined by calculating

CC homology between murine and human antibody surfaces. In order to test

CC the resurfacing approach of the invention, three humanisation experiments

CC were set up: (1) traditional loop grafting; (2) resurfacing approach

CC using most similar chain; and (3) resurfacing approach using human

CC sequences with most similar surface residues. The antibody used was

CC the murine anti-N901 antibody of which the present sequence is the heavy

CC chain variable region. Numbering of the sequence in the specification

CC starts at 118.

XX Sequence 116 AA;

Query Match 50.0%; Score 40; DB 15; Length 116;

Best Local Similarity 46.2%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAYITPGAFDV 14
:111:111:
DB 46 ewwayisgsftl 58

RESULT 14

ARS2065
ID ARS2065 standard; Protein; 117 AA.

AC AAR52065;

DT 11-OCT-1996 (first entry)

DE Heavy chain variable region of humanised N901/636005 antibody.

KW antibody; humanised; murine; human; heavy chain; light; variable;
framework region; complementarity determining region; reshaping;
modelling; surface residue; modify.

OS Chimeric Homo sapiens.

Key Location/Qualifiers

FT 1..30 /label= framework_region_1

FT /note= "FR 1"

FT Region

FT 31..35 /label= complementarity_determining_region_1

FT /note= "CDR 1"

FT Region

FT 36..49 /note= "FR 2"

FT Region

FT 50..59 /note= "CDR 2"

FT Region

FT 60..98 /note= "FR 3"

FT Region

FT 99..107 /note= "CDR 3"

FT Region

FT 13-APR-1994.

PD 07-SEP-1993; 93EP-0307051.

PF 09-SEP-1992; 92US-0942245.

PR (PEDE/) PEDERSEN J T.

PA (IMMU-) IMMUNOGEN INC.

XX Guld BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;

XX WPI: 1994-120230/15.

DR Method of resurfacing of rodent antibodies to produce humanised

PT antibody forms - for producing non-human antibodies with improved

PT therapeutic efficiency by presenting human surface on V-region

PT Example 1; Fig 4B; 230pp; English.

PS Modification of a rodent antibody (Ab) or fragment by resurfacing in

XX order to produce a humanised rodent Ab can be determined by calculating

CC homology between murine and human Ab antibody surfaces. In order to test

CC the resurfacing approach of the invention, three humanisation

CC experiments were set up: (1) traditional loop grafting; (2) resurfacing

CC approach using most similar chain; and (3) resurfacing approach using

CC human sequences with most similar surface residues. The Ab used was the

CC murine anti-N901 Ab (see AAR52065). The AAC36005 Ab (AAR52065) which has

CC 89 percent homology with anti-N901 Ab was used in the resurfacing

CC approach using most similar chain and the present sequence,

CC N901L/AAG36005, was produced with 103 percent homology to anti-N901 Ab.

```

XX      1
SQ      Sequence      117 AA;

Query Match                      50.0%; Score 40; DB 15; Length 117;
Best Local Similarity 46.2%; Pred. No. 15;
Matches      6; Conservative      4; Mismatches      3; Indels      0; Gaps      0;

OY      2 KWEAVITPGAFDV 14
        :|111:1:1:
Db       46 ewayisgsfti 58

RESULT  15
AAR52067
ID      AAR52067 standard; Protein: 117 AA.
XX
AC      AAR52067;
XX
DT      11-OCT-1996 (first entry)
XX
DE      Heavy chain variable region of humanised N901/PL0123 antibody.
XX
KW      antibody; humanised; murine; human; heavy chain; light; variable;
KW      framework region; complementarity determining region; reshaping;
KW      modelling; surface residue; modify.
XX
OS      Chimeric Homo sapiens.
XX
FH      Key
FH      Region
FT      1..30
FT      /label=framework_region_1
FT      /note="FR 1"
FT      Region
FT      31..35
FT      /label=complementarity_determining_region_1
FT      /note="CDR 1"
FT      Region
FT      36..49
FT      /note="FR 2"
FT      Region
FT      50..59
FT      /note="CDR 2"
FT      Region
FT      60..98
FT      /note="FR 3"
FT      Region
FT      99..107
FT      /note="CDR 3"
XX
PM      EP592106-A1.
XX
PD      13-APR-1994.
XX
PE      07-SEP-1993; 93EP-0307051.
XX
PR      09-SEP-1992; 92US-0942245.
XX
PA      (PEDE/) PEDERSEN J T.
XX      (IMMU-) IMMUNOGEN INC.
XX
PI      Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
XX
DR      WPI; 1994-120230/15.
XX
PT      Method of resurfacing of rodent antibodies to produce humanised
PT      antibody forms - for producing non-human antibodies with improved
PT      therapeutic efficiency by presenting human surface on V-region
XX
PS      Example 1; Fig 4B; 230pp; English.
XX
CC      Modification of a rodent antibody (Ab) or fragment by resurfacing in
CC      order to produce a humanised rodent Ab can be determined by calculating
CC      homology between murine and human Ab antibody surfaces. In order to test
CC      the resurfacing approach of the invention, three humanisation
CC      experiments were set up: (1) traditional loop grafting; (2) resurfacing
CC      approach using most similar chain; and (3) resurfacing approach using

```

```

CC      human sequences with most similar surface residues. The Ab used was the
CC      murine anti-N901 Ab (see AAR52061). The PL0123 Ab (AAR52066) which has 74
CC      percent homology with anti-N901 Ab was used in the resurfacing approach
CC      using most similar surface residues and the present sequence,
CC      N901i/PL0123, was produced with 110 percent homology to anti-N901 Ab.
CC      Sequence numbering starts at 118 in the specification.
XX
SQ      Sequence      117 AA;

```

```

Query Match                      50.0%; Score 40; DB 15; Length 117;
Best Local Similarity 46.2%; Pred. No. 15;
Matches      6; Conservative      4; Mismatches      3; Indels      0; Gaps      0;

OY      2 KWEAVITPGAFDV 14
        :|111:1:1:
Db       46 ewayisgsfti 58

```

Search completed: November 19, 2001, 08:12:35
 Job time: 88 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:14 ; Search time 526.89 Seconds
(without alignments)
7.378 Million cell updates/sec

Title: US-09-610-118-63
Perfect score: 80
Sequence: 1 DKWEAYITPGAFDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main.*
1: /cgn2_6/ptodata/2/paa/US008_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US060_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	14	1	PCT-US00-18152-63
2	80	100.0	14	20	US-09-610-118-63
3	80	100.0	14	22	US-09-833-312-63
4	47	58.8	111	16	US-09-213-103-10
5	47	58.8	323	14	US-09-049-022-3
6	47	58.8	325	1	PCT-US98-27049-5
7	47	58.8	325	1	PCT-US99-30452-3
8	47	58.8	325	15	US-09-150-857-10
9	47	58.8	325	15	US-09-150-857-10

10	47	58.8	325	15	US-09-195-896-10	Sequence 10, Appl
11	47	58.8	325	16	US-09-213-103-5	Sequence 5, Appl1
12	47	58.8	325	16	US-09-213-364-5	Sequence 5, Appl1
13	47	58.8	325	16	US-09-284-320-90	Sequence 90, Appl
14	47	58.8	325	16	US-09-292-228-10	Sequence 10, Appl
15	47	58.8	325	21	US-09-717-778-9	Sequence 9, Appl1
16	47	58.8	325	21	US-09-739-451-5	Sequence 5, Appl1
17	47	58.8	325	22	US-09-804-006-10	Sequence 10, Appl
18	47	58.8	325	22	US-09-804-357-10	Sequence 10, Appl
19	47	58.8	325	24	US-60-161-932-2314	Sequence 2314, Ap
20	47	58.8	325	24	US-60-167-217-10139	Sequence 10139, A
21	47	58.8	325	24	US-60-191-637-10148	Sequence 10148, A
22	46	57.5	383	19	US-09-583-110-42857	Sequence 4287, Ap
23	45	56.2	831	1	PCT-US01-08631-38846	Sequence 38846, A
24	44	55.0	169	1	PCT-US97-02318-304	Sequence 304, App
25	44	55.0	169	13	US-08-903-470-304	Sequence 304, App
26	44	55.0	228	21	US-09-758-472-8370	Sequence 8370, Ap
27	44	55.0	354	24	US-60-167-217-9548	Sequence 9548, Ap
28	44	55.0	354	24	US-60-173-464-7681	Sequence 7681, Ap
29	44	55.0	354	24	US-60-191-637-9527	Sequence 9527, Ap
30	44	55.0	354	24	US-60-191-681-7424	Sequence 7424, Ap
31	44	55.0	867	21	US-09-711-164-357	Sequence 357, App
32	44	55.0	867	24	US-60-164-415-357	Sequence 357, App
33	43	53.8	213	24	US-60-312-544-8228	Sequence 8228, Ap
34	43	53.8	424	1	PCT-US01-08631-60293	Sequence 60293, A
35	43	53.8	856	16	US-09-252-691-7158	Sequence 7158, Ap
36	43	53.8	856	16	US-09-252-691C-7158	Sequence 7158, Ap
37	43	53.8	906	1	PCT-US01-08631-48680	Sequence 48680, A
38	43	53.8	1090	1	PCT-US01-08631-48053	Sequence 48053, A
39	43	53.8	1090	1	PCT-US01-08631-48690	Sequence 48690, A
40	42	52.5	838	18	US-09-489-039A-9864	Sequence 9864, Ap
41	42	52.5	869	18	US-09-489-039A-8561	Sequence 8561, Ap
42	41	51.2	149	24	US-60-178-307-1964	Sequence 1964, Ap
43	41	51.2	185	24	US-60-178-307-1964	Sequence 1964, Ap
44	41	51.2	205	1	PCT-US01-08656-7196	Sequence 7196, Ap
45	41	51.2	280	24	US-60-215-161-7865	Sequence 7865, Ap

ALIGNMENTS

RESULT 1
PCT-US00-18152-63
Sequence 63, Application PC/TUS0018152
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-211-228
CURRENT APPLICATION NUMBER: PCT/US00/18152
CURRENT FILING DATE: 2000-06-30
EARLIER APPLICATION NUMBER: 09/503,387
EARLIER FILING DATE: 2/14/00
EARLIER APPLICATION NUMBER: 09/454,824
EARLIER FILING DATE: 12/6/99
EARLIER APPLICATION NUMBER: 09/345,468
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 63
LENGTH: 14
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-18152-63

Query Match 100.0%; Score 80; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAFDV 14
DB 1 DKWEAYITPGAFDV 14

```
RESULT 2
US-09-610-118-63
; Sequence 63, Application US/09610118
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villerval, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; APPLICANT: Gill, D.
; APPLICANT: Olan, M.
; APPLICANT: Kingsbury, G.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-211
; CURRENT APPLICATION NUMBER: US/09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2/14/00
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 12/6/99
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 6/30/99
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-610-118-63

Query Match          100.0%; Score 80; DB 20; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFDV 14
DB 1 DKWEAYITPGAFDV 14

RESULT 3
US-09-832-312-63
; Sequence 63, Application US/09832312
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-63

Query Match          100.0%; Score 80; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFDV 14
DB 1 DKWEAYITPGAFDV 14
```

```
RESULT 4
US-09-213-103-10
; Sequence 10, Application US/09213103
; GENERAL INFORMATION:
; APPLICANT: Egan, Sean
; TITLE OF INVENTION: Mammalian Brainlac Genes
; FILE REFERENCE: 51609
; CURRENT APPLICATION NUMBER: US/09/213,103
; PRIOR FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: CA 2225126
; PRIOR FILING DATE: 1997-12-17
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-213-103-10

Query Match          58.8%; Score 47; DB 16; Length 111;
Best Local Similarity 58.3%; Pred. No. 10;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAF 12
DB 72 DRWPEYVTAGAF 83

RESULT 5
US-09-049-022-3
; Sequence 3, Application US/09049022
; GENERAL INFORMATION:
; APPLICANT: SOPEET, DANIEL R.
; APPLICANT: RUBEN, STEVEN M.
; TITLE OF INVENTION: CARDIAC AND PANCREATIC PROTEIN AND GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/049,022
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/042,855
; FILING DATE: 28-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0620001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 323 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-049-022-3
```

Query Match 58.8%; Score 47; DB 14; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAF 12
1:1 1:1 1:1 1:1
Db 229 DRMPYVTAGAF 239

RESULT 6
PCT-US98-27049-5
; Sequence 5; Application PC/TUS9827049
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Dendriac and Brainiac-3
; FILE REFERENCE: PF464.PCT
; CURRENT APPLICATION NUMBER: PCT/US98/27049
; EARLIER FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/068,006
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/077,687
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/108,928
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US98-27049-5

Query Match 58.8%; Score 47; DB 1; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAF 12
1:1 1:1 1:1 1:1
Db 229 DRMPYVTAGAF 240

RESULT 7
PCT-US99-30452-3
; Sequence 3; Application PC/TUS9930452
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Brainiac-5
; FILE REFERENCE: PF503.PCT
; CURRENT APPLICATION NUMBER: PCT/US99/30452
; EARLIER FILING DATE: 1999-12-20
; EARLIER APPLICATION NUMBER: 60/113,804
; EARLIER FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
PCT-US99-30452-3

Query Match 58.8%; Score 47; DB 1; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAF 12
1:1 1:1 1:1 1:1
Db 229 DRMPYVTAGAF 240

RESULT 8
US-09-150-857-10

; Sequence 10; Application US/09150857
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jiahong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/093001
; CURRENT APPLICATION NUMBER: US/09/150,857
; CURRENT FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-150-857-10

Query Match 58.8%; Score 47; DB 15; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAF 12
1:1 1:1 1:1 1:1
Db 229 DRMPYVTAGAF 240

RESULT 9
US-09-150-857-10
; Sequence 10; Application US/09150857A
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jiahong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/093001
; CURRENT APPLICATION NUMBER: US/09/150,857A
; CURRENT FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-150-857-10

Query Match 58.8%; Score 47; DB 15; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAF 12
1:1 1:1 1:1 1:1
Db 229 DRMPYVTAGAF 240

RESULT 10
US-09-195-896-10
; Sequence 10; Application US/09195896
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jiahong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/109001
; CURRENT APPLICATION NUMBER: US/09/195,896
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 60/108,379
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 09/150,857
; EARLIER FILING DATE: 1998-09-10

; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-193-896-10

Query Match 58.8%; Score 47; DB 15; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
|:| |:| |||
DB 229 DRMPYVTAGAF 240

RESULT 11
US-09-213-103-5
; Sequence 5, Application US/09213103
; GENERAL INFORMATION:
; APPLICANT: Egan, Sean
; TITLE OF INVENTION: Mammalian Brainiac Genes
; FILE REFERENCE: 51609
; CURRENT APPLICATION NUMBER: US/09/213,103
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: CA 2225126
; EARLIER FILING DATE: 1997-12-17
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-213-103-5

Query Match 58.8%; Score 47; DB 16; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
|:| |:| |||
DB 229 DRMPYVTAGAF 240

RESULT 12
US-09-213-364-5
; Sequence 5, Application US/09213364
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Dendriac and Brainiac-3
; FILE REFERENCE: PF464
; CURRENT APPLICATION NUMBER: US/09/213,364
; CURRENT FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/068,006
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/077,687
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/108,928
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-364-5

Query Match 58.8%; Score 47; DB 16; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 DKWEAYITPGAF 12
|:| |:| |||
DB 229 DRMPYVTAGAF 240

RESULT 13
US-09-284-320-90
; Sequence 90, Application US/09284320
; GENERAL INFORMATION:
; APPLICANT: Kato, Seishi et al.
; TITLE OF INVENTION: HUMAN PROTEINS HAVING TRANSMEMBRANE DOMAINS AND DNAs
; FILE REFERENCE: GIN-6705CPUS
; CURRENT APPLICATION NUMBER: US/09/284,320
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: JP 8-301429
; PRIOR FILING DATE: 1996-11-13
; PRIOR APPLICATION NUMBER: PCT/JP97/04056
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila sp.
US-09-284-320-90

Query Match 58.8%; Score 47; DB 16; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
|:| |:| |||
DB 229 DRMPYVTAGAF 240

RESULT 14
US-09-292-228-10
; Sequence 10, Application US/09292228
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Zhou, Jiahong
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334/126001
; CURRENT APPLICATION NUMBER: US/09/292,228
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 09/195,896
; EARLIER FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 60/108,379
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 09/150,857
; EARLIER FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-292-228-10

Query Match 58.8%; Score 47; DB 16; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
|:| |:| |||
DB 229 DRMPYVTAGAF 240

```

RESULT 15
US-09-717-778-9
; Sequence 9, Application US/09717778
; GENERAL INFORMATION:
; APPLICANT: White, David W.
; APPLICANT: Zhou, Jiaqiong
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Stricker-Krongrad, Alain
; APPLICANT: Clausen, Henrik
; TITLE OF INVENTION: LEPTIN INDUCED GENES
; FILE REFERENCE: 07334-318001
; CURRENT APPLICATION NUMBER: US/09/717,778
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 09/292,228
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 09/195,896
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: US 09/150,857
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: US 60/106,378
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-717-778-9

```

```

Query Match 58.8%; Score 47; DB 21; Length 325;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
   | | | | | | | |
Db 229 DRMPPTYTAGAF 240

```

Search completed: November 19, 2001, 08:23:15
 Job time: 728 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:31 ; Search time 11.44 seconds
(without alignments)
6.518 Million cell updates/sec

Title: US-09-610-118-63

Sequence: 1 DKMEAYITPGAFDV 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 25229 seqs, 5326477 residues

Total number of hits satisfying chosen parameters: 25229

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCN_NEW.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	41	51.2	280 5	US-09-897-516-7865 Sequence 7865, Ap
2	40	50.0	221 5	US-09-976-594-664 Sequence 664, App
3	36	45.0	209 5	US-09-815-242-5452 Sequence 5452, Ap
4	36	45.0	209 5	US-09-815-242-12382 Sequence 12382, A
5	36	45.0	209 5	US-09-815-242-12980 Sequence 12980, A
6	36	45.0	794 5	US-09-815-242-5697 Sequence 5697, A
7	36	45.0	802 5	US-09-815-242-12668 Sequence 12668, A
8	35.5	44.4	329 5	US-09-897-516-6199 Sequence 6199, Ap
9	35.5	44.4	944 5	US-09-815-242-11391 Sequence 11391, A
10	35	43.8	359 5	US-09-897-516-5275 Sequence 5275, Ap
11	35	43.8	390 5	US-09-815-242-5284 Sequence 5284, A
12	35	43.8	398 5	US-09-815-242-12599 Sequence 12599, A
13	35	43.8	1049 5	US-09-815-242-10732 Sequence 10732, A
14	35	43.8	244 5	US-09-815-242-10119 Sequence 10119, A
15	34	42.5	347 5	US-09-815-242-10035 Sequence 10035, A
16	34	42.5	358 5	US-09-815-242-12939 Sequence 12939, A
17	34	42.5	359 5	US-09-815-242-5116 Sequence 5116, Ap
18	34	42.5	448 5	US-09-815-242-5806 Sequence 5806, Ap
19	34	42.5	451 5	US-09-815-242-13127 Sequence 13127, A
20	34	42.5	663 5	US-09-815-242-14080 Sequence 14080, A
21	34	42.5	848 5	US-09-897-516-5488 Sequence 5488, Ap
22	34	42.5	1527 5	US-09-897-516-4688 Sequence 4688, Ap
23	33	41.2	19 5	US-09-969-730-226 Sequence 226, App
24	33	41.2	63 5	US-09-646-673A-95 Sequence 95, Appl
25	33	41.2	146 5	US-09-829-631A-11 Sequence 11, Appl
26	33	41.2	210 5	US-09-969-730-115 Sequence 115, App
27	33	41.2	407 5	US-09-815-242-11965 Sequence 11965, A

28	33	41.2	437 5	US-09-829-631A-8 Sequence 8, Appl
29	33	41.2	439 5	US-09-829-631A-13 Sequence 13, Appl
30	33	41.2	440 6	US-60-329-926-10 Sequence 10, Appl
31	33	41.2	462 5	US-09-815-242-5562 Sequence 5562, Ap
32	33	41.2	463 5	US-09-815-242-12572 Sequence 12572, A
33	33	41.2	463 5	US-09-815-242-12754 Sequence 12754, A
34	33	41.2	491 5	US-09-815-242-11274 Sequence 11274, A
35	33	41.2	507 5	US-09-966-707-81 Sequence 81, Appl
36	33	41.2	588 5	US-09-897-516-7344 Sequence 7344, Ap
37	33	41.2	678 5	US-09-897-516-7356 Sequence 7356, Ap
38	33	41.2	937 5	US-09-964-899-51 Sequence 51, Appl
39	33	41.2	1003 5	US-09-370-861A-6 Sequence 6, Appl
40	32.5	40.6	352 5	US-09-761-196A-2 Sequence 2, Appl
41	32	40.0	199 5	US-09-897-516-6189 Sequence 6189, Ap
42	32	40.0	259 5	US-09-897-516-4868 Sequence 4868, Ap
43	32	40.0	261 5	US-09-897-516-4859 Sequence 4859, Ap
44	32	40.0	347 5	US-09-815-242-13914 Sequence 13914, A
45	32	40.0	416 5	US-09-897-516-7213 Sequence 7213, Ap

ALIGNMENTS

```
RESULT 1
US-09-897-516-7865
; Sequence 7865, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hankle, Gregory J.
; APPLICANT: Hoesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Steven M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 7865
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-7865

Query Match          51.2%  Score 41;  DB 5;  Length 280;
Best Local Similarity 85.7%  Pred. No. 1.5;
Matches 6;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

OY 3 WEAYITP 9
   |||||
Db 102 WEAYITP 108

RESULT 2
US-09-976-594-664
; Sequence 664, Application US/09976594
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchdinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROID
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 664
```

```
;; LENGTH: 221
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; OTHER INFORMATION: Incyte ID No: 2636759CD1
US-09-976-594-664
```

```
Query Match      50.0%; Score 40; DB 5; Length 221;
Best Local Similarity 63.6%; Pred. No. 1.7;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy 4 EAYTPGAFD 14
Db 31 EDYCTPGAFEL 41
```

```
RESULT 3
US-09-815-242-5452
; Sequence 5452, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; ORGANISM: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5452
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5452
```

```
Query Match      45.0%; Score 36; DB 5; Length 209;
Best Local Similarity 70.0%; Pred. No. 8.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy 4 EAYTPGAFD 13
Db 191 KAYTPGLGD 200
```

```
RESULT 4
US-09-815-242-12382
; Sequence 12382, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
```

```
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; ORGANISM: Prokaryotes
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12382
;; LENGTH: 209
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-12382
```

```
Query Match      45.0%; Score 36; DB 5; Length 209;
Best Local Similarity 70.0%; Pred. No. 8.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy 4 EAYTPGAFD 13
Db 191 KAYTPGLGD 200
```

```
RESULT 5
US-09-815-242-12980
; Sequence 12980, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; ORGANISM: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
```

```

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12980
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12980

```

```

Query Match      45.0%; Score 36; DB 5; Length 209;
Best Local Similarity 70.0%; Pred. NO. 8.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 4 WEAYTPGAFD 13
    :|||||
Db 191 KAYITPGIGD 200

```

```

RESULT 6
US-09-815-242-5697
; Sequence 5697, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5697
; LENGTH: 794
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-5697

```

```

Query Match      45.0%; Score 36; DB 5; Length 794;
Best Local Similarity 45.5%; Pred. NO. 35;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 3 WEAYTPGAFD 13
    :|||||
Db 427 WITLVTPGTFE 437

```

```

RESULT 7
US-09-815-242-12668
; Sequence 12668, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert

```

```

; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12668
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12668

```

```

Query Match      45.0%; Score 36; DB 5; Length 802;
Best Local Similarity 45.5%; Pred. NO. 35;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 3 WEAYTPGAFD 13
    :|||||
Db 427 WITLVTPGTFE 437

```

```

RESULT 8
US-09-897-516-6199
; Sequence 6199, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkley, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6199
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
; US-09-897-516-6199

```

```

Query Match      44.4%; Score 35.5; DB 5; Length 329;
Best Local Similarity 46.7%; Pred. NO. 16;
Matches 7; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```

OY 1 DKW--EAYITPGAF 12
111 111111
Db 163 DKMTQPEHFVSSGAF 177

RESULT 9

US-09-815-242-11391
; Sequence 11391, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11391
; LENGTH: 944
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11391

Query Match 44.4% Score 35.5; DB 5; Length 944;
Best Local Similarity 63.6% Pred. No. 51;

Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

OY 1 DKWEAYIT-TPG 10
111 111111
Db 491 DKWVSFIDTPG 501

RESULT 10

US-09-897-516-5275
; Sequence 5275, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Cordin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malivar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spilodonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30

; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 5275
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-5275

Query Match 43.8% Score 35; DB 5; Length 359;
Best Local Similarity 58.3% Pred. No. 22;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 KWEAYITPGAFD 13
111 111111
Db 179 KFEAYITKGAFFD 190

RESULT 11

US-09-815-242-5284
; Sequence 5284, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5284
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5284

Query Match 43.8% Score 35; DB 5; Length 390;
Best Local Similarity 53.8% Pred. No. 24;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
111 111111
Db 65 DEVEAYITKVEFD 77

RESULT 12

US-09-815-242-12599
; Sequence 12599, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.

```
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12599
LENGTH: 398
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12599
```

```
Query Match      43.8%  Score 35;  DB 5;  Length 398;
Best Local Similarity 53.8%  Pred. No. 24;
Matches 7;  Conservative 2;  Mismatches 4;  Indels 0;  Gaps 0;
```

```
OY      1  DKWEAYITPGAFD 13
      1 11111111
DB      72  DEVEAYVTKVEFD 84
```

```
RESULT 13
US-09-815-242-10732
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
```

```
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10732
LENGTH: 1049
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10732
```

```
Query Match      43.8%  Score 35;  DB 5;  Length 1049;
Best Local Similarity 46.2%  Pred. No. 70;
Matches 6;  Conservative 3;  Mismatches 4;  Indels 0;  Gaps 0;
```

```
OY      2  KWEAYITPGAFDV 14
      1111111111
DB      538  KKESTYLDPAVFEV 550
```

```
RESULT 14
US-09-815-242-10119
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10119
LENGTH: 244
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10119
```

```
Query Match      42.5%  Score 34;  DB 5;  Length 244;
Best Local Similarity 50.0%  Pred. No. 21;
Matches 6;  Conservative 1;  Mismatches 5;  Indels 0;  Gaps 0;
```

```
OY      1  DKWEAYITPGAF 12
      11111111
DB      38  DMQNAVATPGGY 49
```

```
RESULT 15
US-09-815-242-10035
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
```

```
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10035
; LENGTH: 347
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-815-242-10035
```

```
Query Match          42.5%: Score 34: DB 5; Length 347;
Best Local Similarity 35.7%: Pred No. 31;
Matches 5: Conservative 5: Mismatches 4: Indels 0; Gaps 0;
QY 1 DKWEAYITPGARFV 14
   ::||:|:|:|
DB 82 EEWQAFINNSADV 95
```

Search completed: November 19, 2001, 08:23:31
Job time: 743 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM, protein - protein search, using sw model

Run on: November 19, 2001, 08:13:24 ; Search time 43.5 Seconds
(without alignments)
7.242 Million cell updates/sec

Title: US-09-610-118-63
Perfect score: 80
Sequence: 1 DKWEAYITPGAFDV 14

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/CTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	58.8	325	2	US-09-055-097-5
2	41	51.2	363	1	US-08-488-961-4
3	41	51.2	363	4	US-08-973-297-4
4	41	51.2	363	5	PCR-US96-06511-4
5	41	51.2	411	4	US-09-070-356-6
6	40	50.0	117	1	US-07-942-245-32
7	40	50.0	117	1	US-07-942-245-36
8	40	50.0	117	1	US-07-942-245-38
9	39	48.8	596	1	US-08-565-386-11
10	39	48.8	1045	1	US-08-452-083-2
11	38	47.5	117	2	US-08-822-028-2
12	38	47.5	117	2	US-08-479-285-2
13	38	47.5	133	2	US-08-822-028-6
14	38	47.5	133	2	US-08-822-028-30
15	38	47.5	133	4	US-08-479-285-6
16	38	47.5	133	4	US-08-479-285-30
17	38	47.5	133	6	5219996-17
18	38	47.5	134	2	US-08-822-028-10
19	38	47.5	134	4	US-08-479-285-10
20	38	47.5	150	2	US-08-400-115-2
21	38	47.5	271	2	US-08-400-115-4
22	38	47.5	288	3	US-08-995-280C-2
23	38	47.5	288	3	US-09-215-042-2
24	38	47.5	288	3	US-09-335-409-5
25	37	46.2	363	4	US-08-488-961-7
26	37	46.2	363	4	US-08-973-297-7
27	37	46.2	363	5	PCR-US96-06511-7

28	37	46.2	623	1	US-08-332-838-2	Sequence 2, Appl
29	36	45.0	480	2	US-08-468-812-5	Sequence 5, Appl
30	36	45.0	480	2	US-08-590-563-5	Sequence 5, Appl
31	36	45.0	492	2	US-08-468-812-4	Sequence 4, Appl
32	36	45.0	492	2	US-08-468-812-7	Sequence 7, Appl
33	36	45.0	492	4	US-08-590-563-4	Sequence 4, Appl
34	36	45.0	492	4	US-08-590-563-7	Sequence 7, Appl
35	36	45.0	590	2	US-08-756-317-12	Sequence 12, Appl
36	36	45.0	1024	4	US-09-091-117-5	Sequence 5, Appl
37	36	44.4	46	4	US-08-928-213B-52	Sequence 52, Appl
38	35.5	44.4	219	4	US-08-928-213B-60	Sequence 60, Appl
39	35	43.8	38	1	US-08-176-500-118	Sequence 118, App
40	35	43.8	38	1	US-08-471-052A-118	Sequence 118, App
41	35	43.8	38	1	US-08-189-331-118	Sequence 118, App
42	35	43.8	38	2	US-08-471-939-118	Sequence 118, App
43	35	43.8	38	2	US-08-471-800-118	Sequence 118, App
44	35	43.8	38	2	US-08-471-068-118	Sequence 118, App
45	35	43.8	77	4	US-09-100-802-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-055-097-5
; Sequence 5, Application US/09055097
; Patent No. 5955282
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,097
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerione, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0490 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1150971
; US-09-055-097-5

Query Match 58.88; Score 47; DB 2; Length 325;
Best Local Similarity 58.38; Pred. No. 2.6;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 12
DB 229 DKWEAYITPGAF 240

RESULT 2

US-08-488-961-4
Sequence 4, Application US/08488961
Patent No. 5606042
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
APPLICANT: Walker, John C.
TITLE OF INVENTION: Glycine and Phaseolus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Relsing, Ehlington, Barnard & Perry
STREET: P.O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
ZIP: 48099-4390
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,961
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-320 (UMO)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-488-961-4

Query Match 51.28; Score 41; DB 1; Length 363;
Best Local Similarity 46.28; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 13
DB 201 DKWASYAGPGWN 213

RESULT 3

US-08-973-297-4
Sequence 4, Application US/08973297
Patent No. 6184017
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
APPLICANT: Walker, John C.
TITLE OF INVENTION: Glycine and Phaseolus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6184017thwestern Hwy., Suite 410

CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,297
FILING DATE:
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994,00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5050
TELEFAX: (810) 539-5055

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-973-297-4

Query Match 51.28; Score 41; DB 4; Length 363;
Best Local Similarity 46.28; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DKWEAYITPGAF 13
DB 201 DKWASYAGPGWN 213

RESULT 4

PCT-US96-06511-4
Sequence 4, Application PC/TUS9606511
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
APPLICANT: Walker, John C.
TITLE OF INVENTION: Glycine and Phaseolus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 Northwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06511
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994,00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5050
TELEFAX: (810) 539-5055
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US96-06511-4

Query Match 51.2%; Score 41; DB 5; Length 363;
Best Local Similarity 46.2%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
Db 201 DKWASYAGPGMN 213

RESULT 5
US-09-070-356-6
Sequence 6, Application US/09070356
Patent No. 6228631
GENERAL INFORMATION:
APPLICANT: Alex Zhu
TITLE OF INVENTION: Recombinant a-N-
TITLE OF INVENTION: Acetylglucosaminidase
TITLE OF INVENTION: Enzyme and CDNA Encoding
TITLE OF INVENTION: Said Enzyme
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,356
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/037,248
FILING DATE: March 26, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 63475/12
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581 4766
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: CDNA to mRNA
HYPOTHETICAL: no
ANTI-SENSE: yes
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: guar plant Cyamopsis tetragonoloba
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:

CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: library
POSITION IN GENOME: unknown
CHROMOSOME/SEGMENT:
MAP POSITION:

UNITS:
FEATURE:
NAME/KEY: guar a-galactosidase
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Overbeek et al
TITLE: Cloning and Nucleotide Sequence of
Patent No. 6228631
TITLE: the a-Galactosidase CDNA From
TITLE: Cyamopsis tetragonoloba (guar)
JOURNAL: Plant Molecular Biology
VOLUME: 13
PAGES: 541-550
DATE: 1989
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-09-070-356-6

Query Match 51.2%; Score 41; DB 4; Length 411;
Best Local Similarity 46.2%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DKWEAYITPGAFD 13
Db 248 DKWASYAGPGMN 260

RESULT 6
US-07-942-245-32
Sequence 32, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: in house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-32

Query Match 50.0%; Score 40; DB 1; Length 117;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 2 KWEAYITPGADV 14
Db 46 EWVAYISSGSFTI 58

RESULT 7
US-07-942-245-36
Sequence 36, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrie, Mlon, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7860
TELEFAX: (202) 293-7060
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-36

Query Match 50.0%; Score 40; DB 1; Length 117;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 2 KWEAYITPGADV 14
Db 46 EWVAYISSGSFTI 58

RESULT 8
US-07-942-245-38
Sequence 38, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.

APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrie, Mlon, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7860
TELEFAX: (202) 293-7060
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-38

Query Match 50.0%; Score 40; DB 1; Length 117;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
Db 46 EWVAYISSGSFTI 58

RESULT 9
US-08-565-386-11
Sequence 11, Application US/08565386
Patent No. 5741697
GENERAL INFORMATION:
APPLICANT: Bayoll, Patrick M.
TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Roches
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/565,386
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timlan, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 176/60040
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636

TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 596 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-565-386-11

Query Match 48.8%; Score 39; DB 1; Length 596;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 DKWEAYITPG 10
| 11:111 |
Db 94 DNWSEFITG 103

RESULT 10
; US-08-452-083-2
; Sequence 2, Application US/08452083
; Patent No. 5756327
; GENERAL INFORMATION:
; APPLICANT: Sassanfar, Mandana
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL ISOLEUCYL-tRNA
; TITLE OF INVENTION: SYNTHETASE GENES, TESTER STRAINS AND ASSAYS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452.083
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/305,765
; FILING DATE: 13-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPT94-08B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1045 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-452-083-2

Query Match 48.8%; Score 39; DB 1; Length 1045;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 DKWEAYIT 8
| 1:1111:1 |
Db 144 DEMQAYVT 151

RESULT 11
US-08-822-028-2
; Sequence 2, Application US/08822028
; Patent No. 5993813
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHOLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; TITLE OF INVENTION: MODIFIED ANTIBODIES FOR CANCER TREATMENT
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DUANE C ULMER
; STREET: P.O. BOX 1967
; CITY: MIDLAND
; STATE: MICHIGAN
; COUNTRY: USA
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822.028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/040,687
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-37,075C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-822-028-2

Query Match 47.5%; Score 38; DB 2; Length 117;
Best Local Similarity 46.2%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 KWEAYITPGADV 14
| 1111111 |
Db 65 EWIGYISPGNGDI 77

RESULT 12
US-08-479-285-2
; Sequence 2, Application US/08479285
; Patent No. 6207815
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHOLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; TITLE OF INVENTION: MODIFIED ANTIBODIES FOR CANCER TREATMENT
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:

ADDRESSEE: DUANE C ULMER
STREET: P.O. BOX 1967
CITY: MIDLAND
STATE: MICHIGAN
COUNTRY: USA
ZIP: 48641-1967
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,285
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/040687
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: ULMER, DUANE C
REGISTRATION NUMBER: 34,941
REFERENCE/DOCKET NUMBER: C-37,075C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 636-8104
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-479-285-2

Query Match 47.5% Score 38: DB 4; Length 117:
Best Local Similarity 46.2% Pred. No. 26;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 KWEAYITPGAFDV 14
: 1 11:11 1:
Db 65 EMIGYISPGNDI 77

RESULT 13
US-08-822-028-6
Sequence 6, Application US/08822028
Patent No. 5993813
GENERAL INFORMATION:
APPLICANT: MEZES, PETER S
APPLICANT: GOURLIE, BRIAN B
APPLICANT: RIXON, MARK W
APPLICANT: ANDERSON, WH KERR
APPLICANT: KAPLAN, DONALD A
APPLICANT: SCHOLOM, JEFFREY
TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
MODIFIED ANTIBODIES FOR CANCER TREATMENT
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: DUANE C ULMER
STREET: P.O. BOX 1967
CITY: MIDLAND
STATE: MICHIGAN
COUNTRY: USA
ZIP: 48641-1967
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,028
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/040,687
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: ULMER, DUANE C
REGISTRATION NUMBER: 34,941
REFERENCE/DOCKET NUMBER: C-37,075C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 636-8104
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 133 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-028-6

Query Match 47.5% Score 38: DB 2; Length 133:
Best Local Similarity 46.2% Pred. No. 30;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 KWEAYITPGAFDV 14
: 1 11:11 1:
Db 65 EMIGYISPGNDI 77

RESULT 14
US-08-822-028-30
Sequence 30, Application US/08822028
Patent No. 5993813
GENERAL INFORMATION:
APPLICANT: MEZES, PETER S
APPLICANT: GOURLIE, BRIAN B
APPLICANT: RIXON, MARK W
APPLICANT: ANDERSON, WH KERR
APPLICANT: KAPLAN, DONALD A
APPLICANT: SCHOLOM, JEFFREY
TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
MODIFIED ANTIBODIES FOR CANCER TREATMENT
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: DUANE C ULMER
STREET: P.O. BOX 1967
CITY: MIDLAND
STATE: MICHIGAN
COUNTRY: USA
ZIP: 48641-1967
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,028
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/040,687
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: ULMER, DUANE C
REGISTRATION NUMBER: 34,941
REFERENCE/DOCKET NUMBER: C-37,075C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 636-8104
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 133 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-028-30

Query Match 47.5%; Score 38; DB 2; Length 133;
Best Local Similarity 46.2%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
: | | | | | :
Db 65 EWIGYISGNDI 77

RESULT 15

US-08-479-285-6
; Sequence 6, Application US/08479285
; Patent No. 6207815
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHOLIM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; NUMBER OF SEQUENCES: 74
; TITLE OF INVENTION: MODIFIED ANTIBODIES FOR CANCER TREATMENT
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DUANE C ULMER
; STREET: P.O. BOX 1967
; CITY: MIDLAND
; STATE: MICHIGAN
; COUNTRY: USA
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,285
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/040687
; FILING DATE: 31-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-37,075C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-479-285-6

Query Match 47.5%; Score 38; DB 4; Length 133;
Best Local Similarity 46.2%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 KWEAYITPGADV 14
: | | | | | :
Db 65 EWIGYISGNDI 77

Search completed: November 19, 2001, 08:13:25
Job time: 138 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:14:21 ; Search time 48.99 Seconds
(Without alignments)
20.214 Million cell updates/sec

Title: US-09-610-118-64
Perfect score: 62
Sequence: 1 TRSSGSIASNYVQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	83.9	112	1	Ig lambda chain V-
2	52	83.9	131	1	Ig lambda chain V-
3	51	82.3	111	1	Ig lambda chain V-
4	43	69.4	111	1	Ig lambda chain V-
5	40	64.5	778	2	Ig lambda chain V-
6	38	61.3	112	1	Ig lambda chain V-
7	38	61.3	505	1	Ig lambda chain V-
8	38	61.3	556	2	Ig lambda chain V-
9	38	61.3	606	2	Ig lambda chain V-
10	38	61.3	726	2	Ig lambda chain V-
11	38	61.3	726	2	Ig lambda chain V-
12	37	59.7	108	2	Ig kappa chain V r
13	37	59.7	236	2	Ig kappa chain V r
14	37	59.7	1181	2	Ig kappa chain V r
15	36	58.1	140	2	Ig kappa chain V r
16	36	58.1	352	2	Ig kappa chain V r
17	36	58.1	522	2	Ig kappa chain V r
18	36	58.1	539	2	Ig kappa chain V r
19	36	58.1	549	2	Ig kappa chain V r
20	36	58.1	997	2	Ig kappa chain V r
21	35	56.5	131	1	Ig lambda chain pr
22	35	56.5	278	2	Ig lambda chain pr
23	35	56.5	402	2	Ig lambda chain pr
24	35	56.5	503	1	Ig lambda chain pr
25	35	56.5	503	1	Ig lambda chain pr
26	35	56.5	512	1	Ig lambda chain pr
27	35	56.5	537	2	Ig lambda chain pr
28	35	56.5	541	2	Ig lambda chain pr
29	35	56.5	574	2	Ig lambda chain pr

30	35	56.5	1007	2	H81670	conserved hypotet
31	35	56.5	1167	2	T06146	disease resistance
32	34	54.8	47	2	B30607	Ig kappa chain V-I
33	34	54.8	55	2	P00829	E2/NS1 protein (cd
34	34	54.8	72	2	I37232	calctonin gene-re
35	34	54.8	98	2	S36048	Ig lambda chain -
36	34	54.8	124	2	C48234	hypothetical prote
37	34	54.8	127	2	A25864	calctonin gene-re
38	34	54.8	129	2	S78058	Ig lambda chain pr
39	34	54.8	130	2	S78057	Ig lambda chain pr
40	34	54.8	237	2	JC7217	paternally imprinf
41	34	54.8	294	2	J00798	nucleoside-specifi
42	34	54.8	294	2	D85537	hypothetical prote
43	34	54.8	350	2	I46608	MHC PD6-glycoprote
44	34	54.8	367	2	S50621	SSU81 protein - ye
45	34	54.8	396	2	T01049	hypothetical prote

ALIGNMENTS

RESULT 1

L6HU48
Ig lambda chain V-VI region (Nig-48) - human
C:Species: Homo sapiens (man)
C:Date: 28-Feb-1980 #sequence_revision 28-Feb-1980 #text_change 02-Sep-1997
C:Accession: A01991
R:Takahashi, N.; Takayasu, T.; Isobe, T.; Shinoda, T.; Okuyama, T.; Shimizu, A.
J. Biochem. 86, 1523-1535, 1979
A:Title: Comparative study on the structure of the light chains of human immunoglobul
A:Reference number: A01991; MUID:80094390
A:Accession: A01991
A:Molecule type: protein
A:Residues: 1-112 <TRAK>
A:Note: this is the first sequenced V region of lambda chain subgroup VI
C:Comment: This is a Bence Jones protein.
C:Genetics:
A:Gene: GDB:IGLV6
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer
F:15-93/Domain: Immunoglobulin homology <IMM>
F:22-91/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 83.9%; Score 52; DB 1; Length 112;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYVQ 13
DB 23 TRSSGSIASNYVQ 35

RESULT 2

L6HU48
Ig lambda chain precursor V-VI region (EB4) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
C:Accession: A01990
R:Anderson, M.L.M.; Brown, L.; McKenzie, E.; Kellow, J.E.; Young, B.D.
Nucleic Acids Res. 13, 2931-2941, 1985
A:Title: Cloning and sequence analysis of an Ig lambda light chain mRNA expressed in A:Reference number: A01990; MUID:85215660
A:Accession: A01990
A:Molecule type: mRNA
A:Residues: 1-131 <AND>
C:Genetics:
A:Gene: GDB:IGLV6
A:Cross-references: GDB:119342; OMIM:147240

A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (Kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-19/Domain: signal sequence #status predicted <Stc>
F:20-131/Product: Ig lambda chain V-VI region (EB4) #status predicted <MAT>
F:20-41/Region: framework 1
F:34-112/Domain: immunoglobulin homology <IMM>
F:42-54/Region: complementarity-determining 1
F:55-69/Region: framework 2
F:70-76/Region: complementarity-determining 2
F:77-110/Region: framework 3
F:111-118/Region: complementarity-determining 3
F:119-131/Region: framework 4
F:41-110/Disulfide bonds: #status predicted

Query Match 83.9%; Score 52; DB 1; Length 131;
Best Local Similarity 84.6%; Pred. No. 0.015;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYVQ 13
1:|||||
DB 42 TGSSGSIASNYVQ 54

RESULT 3

L6HUR
Ig lambda chain V-VI region (WLT) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
C:Accession: A01989
R:Dwulet, F.E.; Strako, K.; Benson, M.D.
Scand. J. Immunol. 22, 653-660, 1985
A:Title: Amino acid sequence of a lambda VI primary (AL) amyloid protein (WLT).
A:Reference number: A01989; MUID:8612267
A:Accession: A01989
A:Molecule type: protein
A:Residues: 1-111 <DMC>
C:Genetics:
A:Gene: GDB:IGLV4
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (Kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-32/Region: framework 1
F:15-93/Domain: immunoglobulin homology <IMM>
F:23-35/Region: complementarity-determining 1
F:36-50/Region: framework 2
F:51-57/Region: complementarity-determining 2
F:58-91/Region: framework 3
F:92-101/Region: complementarity-determining 3
F:102-111/Region: framework 4
F:22-91/Disulfide bonds: #status predicted

Query Match 82.3%; Score 51; DB 1; Length 111;
Best Local Similarity 84.6%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYVQ 13
1:|||||
DB 23 TGSSGSIASNYVQ 35

RESULT 4

L6HUR
Ig lambda chain V-VI region (SUT) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997

C:Accession: A01988
R:Solomon, A.; Kyle, R.A.; Frangione, B.
In Amyloidosis, Glenner, G.G.; Osserman, E.F.; Benditt, E.P.; Calkins, E.; Cohn, A.S.
A:Title: Light chain variable region subgroups of monoclonal immunoglobulins in amyloidosis
A:Reference number: A01988
A:Accession: A01988
A:Molecule type: protein
A:Residues: 1-111 <SOL>
C:Genetics:
A:Gene: GDB:IGLV4
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (Kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-22/Region: framework 1
F:15-93/Domain: immunoglobulin homology <IMM>
F:23-35/Region: complementarity-determining 1
F:36-50/Region: framework 2
F:51-57/Region: complementarity-determining 2
F:58-91/Region: framework 3
F:92-101/Region: complementarity-determining 3
F:101-111/Region: framework 4
F:22-91/Disulfide bonds: #status predicted

Query Match 69.4%; Score 43; DB 1; Length 111;
Best Local Similarity 69.2%; Pred. No. 0.63;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYVQ 13
1:|||||
DB 23 TRSDGTAGYVQ 35

RESULT 5

T39047
hypothetical protein SPAC6F6.13c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 15-Sep-2000
C:Accession: T39047
R:Gentiles, S.; Churche, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: 221797
A:Accession: T39047
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-778 <GEN>
A:Cross-references: EMBL:298981; PIDN:CA811736.1; GSPDB:GM00066; SPDB:SPAC6F6.13c
A:Experimental source: strain 972h-; cosmid c6F6
C:Genetics:
A:Gene: SPDB:SPAC6F6.13c
A:Map position: 1
C:Superfamily: Candida albicans conserved hypothetical protein PL09050.2

Query Match 64.5%; Score 40; DB 2; Length 778;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 SSGSIASNYVQ 13
1:|||||
DB 39 SKGSTENYVQ 49

RESULT 6

L6HUR
Ig lambda chain V-VI region (AR) - human (tentative sequence)
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 31-Mar-2000
C:Accession: A01987
R:Stetten, K.; Natvig, J.B.; Husby, G.; Juul, J.

A:Gene: CESP:Y51A2D.4
A:Map position: 5
A:Introns: 4/3; 76/1; 144/1; 226/3; 400/2; 446/1; 562/3
C:Superfamily: glucose transport protein

Query Match 61.3%; Score 38; DB 2; Length 606;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYV 11
| : ||||| : |
Db 442 TNSNGSIANNF 452

RESULT 10
E85633
hypothetical protein yccc [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: E85633
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85633
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-726 <STO>
A:Cross-references: GB:AE005174; NID:g12514246; PIDN:AGC55529.1; GSPDB:GN00145; UWGP:213
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yccc

Query Match 61.3%; Score 38; DB 2; Length 726;
Best Local Similarity 61.5%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYV 13
| : ||||| : |
Db 243 TRILNSIANNYIQ 255

RESULT 11
C64839
probable ATPase yccc - Escherichia coli
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 02-Feb-2001
C:Accession: C64839
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: C64839
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-726 <BLAT>
A:Cross-references: GB:AE000200; GB:U00096; NID:g2367111; PIDN:AAC74066.1; PID:g1787216;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yccc
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F:34-50/Domain: transmembrane status predicted <TM>
F:428-444/Domain: transmembrane status predicted <TM>
F:539-546/Region: nucleotide-binding motif A (P-loop)

Query Match 61.3%; Score 38; DB 2; Length 726;
Best Local Similarity 61.5%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYV 13
| : ||||| : |
Db 243 TRILNSIANNYIQ 255

RESULT 12
PS0073
Ig kappa chain V region (38C13.V2) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jan-2000
C:Accession: PS0073
R:Levy, S.; Campbell, M.J.; Levy, R.
J. Exp. Med. 170, 1-13, 1989
A:Title: Functional Immunoglobulin light chain genes are replaced by ongoing rearrang
A:Reference number: A92781; MUID:89310348
A:Accession: PS0073
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-108 <LEV>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:16-91/Domain: Immunoglobulin homology <TM>

Query Match 59.7%; Score 37; DB 2; Length 108;
Best Local Similarity 63.6%; Pred. No. 8.4;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 RSSGSIASNYV 12
| : ||||| : |
Db 24 RASSSISSNYL 34

RESULT 13
E59098
hypothetical protein PX01-61 - Bacillus anthracis virulence plasmid PX01
C:Species: Bacillus anthracis
C>Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 16-Feb-2001
C:Accession: E59098
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Kelm, P.; Koeh
J. Bacteriol. 181, 6509-6515, 1999
A:Title: Sequence and organization of PX01, the large Bacillus anthracis plasmid harb
A:Reference number: A59091; MUID:99445483
A:Accession: E59098
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-236 <OKI>
A:Cross-references: GB:AF065404; NID:g4894216; PIDN:AMD32365.1; PID:g4894277
A:Experimental source: strain Sterne
C:Genetics:
A:Gene: PX01-61
A:Genome: plasmid
C:Superfamily: hypothetical protein PX01-61

Query Match 59.7%; Score 37; DB 2; Length 236;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 RSSGSIASNYV 13
| : ||||| : |
Db 25 KDSGSLNYYE 36

RESULT 14
T30578
myosin IC - slime mold (Dictyostellium discoideum)
C:Species: Dictyostellium discoideum
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 08-Sep-2000
C:Accession: T30578
R:Peterson, M.D.; Novak, K.D.; Reedy, M.C.; Ruman, J.I.; Titus, M.A.
J. Cell Sci. 108, 1093-1103, 1995

A:Title: Molecular genetic analysis of myoC, a Dictyostelium myosin I.
 A:Reference number: Z20872; MUID:95348228
 A:Accession: T30578
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1181 <PEP>
 A:Cross-references: EMBL:L35323; NID:9532123; PID:9532124; PIDN:AAC37427.1
 C:Genetics:
 A:Gene: myoC
 A:Introns: 307/1
 C:Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homolo
 F:18-685/Domain: myosin motor domain homology <MMO>

Query Match 59.7%; Score 37; DB 2; Length 1181;
 Best Local Similarity 53.8%; Pred. No. 95;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 TRSSGSIASNTYQ 13
 : : : : :
 Db 1167 TKQIGMLPSNTYQ 1179

RESULT 15

H86292
 hypothetical protein AAF82154.1 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86292
 R:Neologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: H86292
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-140 <STO>
 A:Cross-references: GB:AE05172; NID:98927663; PIDN:AAF82154.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 58.1%; Score 36; DB 2; Length 140;
 Best Local Similarity 88.9%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRSSGSIAS 9
 : : : : :
 Db 56 TRSDGSIAS 64

Search completed: November 19, 2001, 08:14:22
 Job time: 195 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:35 : Search time 27.32 Seconds
(without alignments)
16,300 Million cell updates/sec

Title: US-09-610-118-64

Perfect score: 62
Sequence: 1 TRSSGSIASNYQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	52	83.9	112	1 LV6B_HUMAN P01722 homo sapien
2	52	83.9	131	1 LV6E_HUMAN P06319 homo sapien
3	51	82.3	111	1 LV6D_HUMAN P06318 homo sapien
4	43	69.4	111	1 LV6C_HUMAN P06317 homo sapien
5	38	61.3	112	1 LV6A_HUMAN P01721 homo sapien
6	38	61.3	526	1 HCK_HUMAN P08631 homo sapien
7	38	61.3	726	1 ERK_ECOLI P38134 escherichia
8	37	59.7	1181	1 MYSC_DICDI P42522 dictyostell
9	36	58.1	1	1 PXN1_XENLA O13297 xenopus lae
10	36	58.1	549	1 CER1_YEAST P50545 rattus norv
11	35	56.5	503	1 HCK_RAT P07948 homo sapien
12	35	56.5	511	1 LVN_HUMAN P08103 mus musculu
13	35	56.5	524	1 HCK_MOUSE P27446 xiphophorus
14	35	56.5	536	1 FYN_XIPHE O04736 mus musculu
15	35	56.5	541	1 YES_MOUSE P87311 schizosacch
16	35	56.5	574	1 YBZA_SCHPO P30880 sus scrofa
17	34	54.8	37	1 CGRP_PIG P10092 homo sapien
18	34	54.8	127	1 CAL2_HUMAN P22786 escherichia
19	34	54.8	294	1 TSX_ECOLI P40073 saccharomyc
20	34	54.8	367	1 SSB1_YEAST P21111 squitrel mo
21	34	54.8	740	1 GAG_SMRVH P55038 synchocyst
22	34	54.8	1556	1 GLRS_SYNY3 P30881 ovils aries
23	33	53.2	37	1 GRP_SHEEP O99050 bombyx mand
24	33	53.2	178	1 FBOH_BOVMA P75866 escherichia
25	33	53.2	182	1 YBBA_ECOLI P13696 homo sapien
26	33	53.2	186	1 PEPB_BOVIN P30086 homo sapien
27	33	53.2	186	1 PEPB_HUMAN P48737 macaca fasc
28	33	53.2	186	1 PEPB_MACFA P70296 mus musculu
29	33	53.2	187	1 PEPB_MOUSE O9xy40 drosophila
30	33	53.2	1	1 CRK_DROME O96483 lycopersico
31	33	53.2	336	1 ACT2_LYCES O80934 arabidopsis
32	33	53.2	337	1 Y230_ARATH P38200 saccharomyc
33	33	53.2	338	1 YBD1_YEAST

34	33	53.2	339	1 ACT3_TOBAC P93373 nicotiana t
35	33	53.2	351	1 REBB_XANCA P55295 xanthomonas
36	33	53.2	368	1 YBHR_ECOLI P75774 escherichia
37	33	53.2	425	1 KDTA_ECOLI P23282 escherichia
38	33	53.2	504	1 GAG_HVIMA P4594 human immun
39	33	53.2	533	1 INV_DEBOC P24133 debaryomyce
40	33	53.2	590	1 SRC2_DROME P08630 drosophila
41	33	53.2	602	1 ELU_MOUSE O08836 mus musculu
42	33	53.2	691	1 HSF_DROME P22813 drosophila
43	33	53.2	1433	1 REST_CHICK O42184 gallus gall
44	33	53.2	1474	1 A2MG_HUMAN P01023 homo sapien
45	33	53.2	1482	1 P2P_HUMAN P20742 homo sapien

ALIGNMENTS

RESULT 1	LV6B_HUMAN	STANDARD;	PRT;	112 AA.
ID	LV6B_HUMAN			
AC	P01722;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG LAMBDA CHAIN V-VI REGION NIG-48.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=80094390; PubMed=118171;			
RA	Takahashi N., Takayasu T., Isoe T., Shinoda T., Okuyama T.,			
RA	Shimizu A.;			
RT	"Comparative study on the structure of the light chains of human			
RT	immunoglobulins. II. Assignment of a new subgroup.";			
RL	J. Biochem. 86:1523-1535(1979).			
CC	-I- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.			
DR	PIR; A01991; L6H048.			
DR	HSSP; P01703; 7FAB.			
DR	InterPro: IPR003006;			
DR	PIfam: PF00047; ig: 1.			
FW	Immunoglobulin V region; Bence-Jones protein.			
KW	NON_TER			
SO	SEQUENCE 112 AA; 12152 MW; CPBB307BC527A384 CRC64;			
Query Match	83.9%;	Score 52;	DB 1;	Length 112;
Best Local Similarity	84.6%;	Pred. No. 0.0033;		
Matches 11;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
OY	1 TRSSGSIASNYQ 13			
DB	23 TRSSGSIASNYQ 35			
RESULT 2	LV6E_HUMAN	STANDARD;	PRT;	131 AA.
ID	LV6E_HUMAN			
AC	P06319;			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG LAMBDA CHAIN V-VI REGION E84 PRECURSOR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215660; PubMed=3923440;			
RA	Anderson M.L.M., Brown L., McKenzie E., Kellow J.E., Young B.D.;			
RT	"Cloning and sequence analysis of an Ig lambda light chain mRNA			

RT expressed in the Burkitt's lymphoma cell line EB4.
 RL Nucleic Acids Res. 13:2931-2941(1985).
 DR PIR: A01990; L6HUB.
 DR HSSP: P01709; 2MCG.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 131 IG LAMBDA CHAIN V-VI REGION EB4.
 FT DOMAIN 20 41 FRAMEWORK 1.
 FT DOMAIN 42 54 COMPLEMENTARITY-DETERMINING 1.
 FT DOMAIN 55 69 FRAMEWORK 2.
 FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING 2.
 FT DOMAIN 110 118 FRAMEWORK 3.
 FT DOMAIN 111 118 COMPLEMENTARITY-DETERMINING 3.
 FT DOMAIN 119 131 FRAMEWORK 4.
 FT DISULFID 41 110 BY SIMILARITY.
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 14147 MW; 02A9179C8C05C2CD CRC64;

Query Match
 Best Local Similarity 83.9%; Score 52; DB 1; Length 131;
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVQ 13
 Db 42 TCSSGSIASNYVQ 54

RESULT 3
 LV6D_HUMAN STANDARD; PRT: 111 AA.
 ID P06318;
 AC 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG LAMBDA CHAIN V-VI REGION WLT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86122667; PubMed=4089539;
 RA Dwyer F.E., Strako K., Benson M.D.;
 RT "Amino acid sequence of a lambda VI primary (AL) amyloid protein
 (WLT)."
 RL Scand. J. Immunol. 22:653-660(1985).
 DR PIR: A01989; L6HULT.
 DR HSSP: P01709; 2MCG.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 22 FRAMEWORK 1.
 FT DOMAIN 23 35 COMPLEMENTARITY-DETERMINING 1.
 FT DOMAIN 36 50 FRAMEWORK 2.
 FT DOMAIN 51 57 COMPLEMENTARITY-DETERMINING 2.
 FT DOMAIN 92 91 FRAMEWORK 3.
 FT DOMAIN 101 101 COMPLEMENTARITY-DETERMINING 3.
 FT DISULFID 102 111 FRAMEWORK 4.
 FT NON_TER 111 111 BY SIMILARITY.
 SQ SEQUENCE 111 AA; 11966 MW; 0C8B2FE37BCE24F CRC64;

Query Match
 Best Local Similarity 82.3%; Score 51; DB 1; Length 111;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVQ 13
 Db 23 TCSSGSIASNYVQ 35

RESULT 4
 LV6C_HUMAN STANDARD; PRT: 111 AA.
 ID P06317;
 AC 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG LAMBDA CHAIN V-VI REGION SUT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RA Solomon A., Kyle R.A., Frangione B.;
 RT "Light chain variable region subgroups of monoclonal immunoglobulins
 in amyloidosis AL."
 RL (In) Glenner G.G., Osseman E.F., Benditt E.P., Calkins E.,
 Cohen A.S., Zucker-Franklin D. (eds.);
 RL Amyloidosis, pp.449-462, Plenum Press, New York (1986).
 DR PIR: A01988; L6HUST.
 DR HSSP: P01703; 7FAB.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 22 FRAMEWORK 1.
 FT DOMAIN 23 35 COMPLEMENTARITY-DETERMINING 1.
 FT DOMAIN 36 50 FRAMEWORK 2.
 FT DOMAIN 51 57 COMPLEMENTARITY-DETERMINING 2.
 FT DOMAIN 92 91 FRAMEWORK 3.
 FT DOMAIN 101 100 COMPLEMENTARITY-DETERMINING 3.
 FT DISULFID 101 111 FRAMEWORK 4.
 FT NON_TER 111 111 BY SIMILARITY.
 SQ SEQUENCE 111 AA; 12247 MW; 0941DD547D983598 CRC64;

Query Match
 Best Local Similarity 69.4%; Score 43; DB 1; Length 111;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVQ 13
 Db 23 TRSDGTAGTYVQ 35

RESULT 5
 LV6A_HUMAN STANDARD; PRT: 112 AA.
 ID P01721;
 AC 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG LAMBDA CHAIN V-VI REGION AR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE (AMYLOID PROTEIN AR).
 RX MEDLINE=82091000; PubMed=6797401;
 RA Sletten K., Natvig J.B., Husby G., Juul J.;
 RT "The complete amino acid sequence of a prototype
 immunoglobulin-lambda light-chain-type amyloid-fibril protein AR."
 RL Biochem. J. 195;561-572(1981).
 CC -1- MISCELLANEOUS: ABOUT HALF OF THE LAMBDA CHAIN C REGION IS MISSING
 FROM THIS PROTEIN
 CC -1- MISCELLANEOUS: THIS PROTEIN WAS ISOLATED FROM THE SPLEEN OF A
 PATIENT WITH AMYLOIDOSIS.
 DR PIR: A01987; L6HUAR.
 DR HSSP: P01709; 2MCG.

DR InterPro: IPR003006; -.
 KW Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region; Amyloid.
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA; 11918 MW; 570BCD9A368EFLFE CRC64;

Query Match 61.3%; Score 38; DB 1; Length 112;
 Best Local Similarity 61.5%; Pred. No. 1.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSGGSINAYVQ 13
 1 1111::11
 Db 23 TSGSGSINSEVQ 35

RESULT 6
 HCK_HUMAN STANDARD; PRT; 526 AA.
 ID HCK_HUMAN
 AC P08631;
 DT 01-AUG-1998 (Rel. 08, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TYROSINE-PROTEIN KINASE HCK (EC 2.7.1.112) (P59-HCK AND P60-HCK)
 DE (HEMOPOIETIC CELL KINASE).
 GN HCK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE OF 22-526 FROM N.A.
 RX MEDLINE=87257942; PubMed=3496523;
 RA Quintrell N., Iabo R., Varmus H., Bishop J.M., Pietenati M.J.,
 le Beau M.M., Diaz M.O., Rowley J.D.;
 "Identification of a human gene (HCK) that encodes a protein-tyrosine
 kinase and is expressed in hemopoietic cells.";
 RT Mol. Cell. Biol. 7:2267-2275(1987).
 RN [2]
 RP SEQUENCE OF 22-526 FROM N.A.
 RX MEDLINE=87257943; PubMed=3453117;
 RA Ziegler S.F., March J.D., Lewis D.B., Perlmutter R.M.;
 "Novel protein-tyrosine kinase gene (hck) preferentially expressed in
 cells of hematopoietic origin.";
 RT Mol. Cell. Biol. 7:2276-2285(1987).
 RN [3]
 RP SEQUENCE OF 1-22 FROM N.A., AND ALTERNATIVE INITIATION.
 RX MEDLINE=91342636; PubMed=1875927;
 RA Lock P., Ralph S., Stanley E., Boulet I., Ramsay R., Dunn A.R.;
 "Two isoforms of murine hck, generated by utilization of alternative
 translational initiation codons, exhibit different patterns of
 subcellular localization.";
 RT Mol. Cell. Biol. 11:4363-4370(1991).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 78-526.
 RX MEDLINE=9717106; PubMed=9024658;
 RA Slicher F., Moarefi I., Kurlyan J.;
 "Crystal structure of the Src family tyrosine kinase Hck.";
 RT Nature 385:602-609(1997).
 RN [5]
 RP STRUCTURE BY NMR OF 78-138.
 RX MEDLINE=98239731; PubMed=9571048;
 RA Horita D.A., Baldisseri D.M., Zhang W., Altieri A.S., Smithgall T.E.,
 Gmeiner W.H., Byrd R.A.;
 "Solution structure of the human Hck SH3 domain and identification of
 its ligand binding site.";
 RT J. Mol. Biol. 278:253-265(1998).
 RN [6]
 RP STRUCTURE BY NMR OF 139-245.
 RX MEDLINE=97263487; PubMed=9109402;
 RA Zhang W., Smithgall T.E., Gmeiner W.H.;
 "Sequential assignment and secondary structure determination for the
 Src homology 2 domain of hematopoietic cellular kinase.";

RL FEBS Lett. 406:131-135(1997).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 81-137.
 RX MEDLINE=98453315; PubMed=9778343;
 RA Arnold S., O'Brien R., Franken P., Strub M.P., Hof F., Dumas C.,
 Ladbury J.E.;
 "RT loop flexibility enhances the specificity of src family SH3
 domains for HIV-1 Nef.";
 RT Biochemistry 37:14683-14691(1998).
 RL CC
 CC -1- FUNCTION: MAY SERVE AS PART OF A SIGNALING PATHWAY COUPLING THE FC
 RECEPTOR TO THE ACTIVATION OF THE RESPIRATORY BURST. MAY ALSO
 CONTRIBUTE TO NEUTROPHIL MIGRATION AND MAY REGULATE THE
 DEGRANULATION PROCESS OF NEUTROPHILS.
 CC CC
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
 PROTEIN TYROSINE PHOSPHATE.
 CC CC
 CC -1- SUBCELLULAR LOCATION: P60-HCK AND P59-HCK ARE ASSOCIATED WITH
 MEMBRANES. P60-HCK IS ALSO CYTOPLASMIC (BY SIMILARITY).
 CC CC
 CC -1- ALTERNATIVE PRODUCTS: THE P60-HCK AND P59-HCK ARE PRODUCED BY THE
 USE OF ALTERNATIVE INITIATION SITES.
 CC CC
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CELLS OF THE
 MYELOID AND B-LYMPHOID LINEAGES.
 CC CC
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC CC
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC CC
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 DOMAIN. BELONGS TO THE SRC SUBFAMILY.
 CC CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 or send an email to license@isb-sib.ch).
 CC CC
 CC EMBL: M16591; AAA52643.1; -;
 CC EMBL: M16592; AAA52644.1; -;
 CC PIR: A27812; TYRHCK.
 CC DR
 CC PDB: 2HCK; 20-AUG-97.
 CC DR
 CC PDB: 3HCK; 15-OCT-97.
 CC DR
 CC PDB: 4HCK; 17-JUN-98.
 CC DR
 CC PDB: 5HCK; 17-JUN-98.
 CC DR
 CC PDB: 1AD5; 15-MAY-97.
 CC DR
 CC PDB: 1BU1; 11-NOV-98.
 CC DR
 CC MIM: 142370; -;
 CC DR
 CC InterPro: IPR000719; -;
 CC DR
 CC InterPro: IPR000980; -;
 CC DR
 CC InterPro: IPR001245; -;
 CC DR
 CC InterPro: IPR001452; -;
 CC DR
 CC Pfam: PF00017; SH2; 1.
 CC DR
 CC Pfam: PF00018; SH3; 1.
 CC DR
 CC Pfam: PF00069; PKINASE; 1.
 CC DR
 CC PRINTS: PR00109; TYRKINASE.
 CC DR
 CC PRINTS: PR00401; SH2DOMAIN.
 CC DR
 CC PRINTS: PR00452; SH3DOMAIN.
 CC DR
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR
 CC PROSITE: PS50001; SH2; 1.
 CC DR
 CC PROSITE: PS50002; SH3; 1.
 CC KW
 KW Transferrase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;
 KW Myristate; SH2 domain; SH3 domain; Alternative initiation;
 KW 3D-structure.
 KW FT
 KW CHAIN 1 526 TYROSINE-PROTEIN KINASE P60-HCK.
 KW FT
 KW CHAIN 22 526 TYROSINE-PROTEIN KINASE P59-HCK.
 KW FT
 KW INT_MET 22 22 FOR P59-HCK.
 KW FT
 KW LIPID 2 2 MYRISTATE (BY SIMILARITY).
 KW FT
 KW LIPID 23 23 MYRISTATE (BY SIMILARITY).
 KW FT
 KW DOMAIN 78 138 SH3.
 KW FT
 KW DOMAIN 144 241 SH2.
 KW FT
 KW DOMAIN 262 515 PROTEIN KINASE.
 KW FT
 KW NP_BIND 268 276 ATP.
 KW FT
 KW BINDING 290 290 ATP.
 KW FT
 KW ACT_SITE 381 381

```

FT MOD_RES 411 411 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 24 24 S -> C (IN AAA52644).
SQ SEQUENCE 526 AA: 59583 MW: 347877AA0A6412B3 CRC64:

Query Match
Best Local Similarity 61.3%; Score 38; DB 1; Length 526;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
Db 122 TRKEGIIPSRYV 133

RESULT 7
ETK_ECOLI STANDARD: PRT: 726 AA.
AC P38134; P75879;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TYROSINE-PROTEIN KINASE ETK (EC 2.7.1.112).
GN ETK.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RA Ilan O.A., Bloch Y., Frankel G., Ullrich H., Geider K., Rosenblum I.;
RT "Protein tyrosine kinases in bacterial pathogens are associated with
virulence and production of exopolysaccharide.";
RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 628-726 FROM N.A.
RC MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [5]
RP IDENTIFICATION.
RC MEDLINE=95075659; PubMed=7984428;
RA Bqodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
bacterial genome.";
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -1- SIMILARITY: SOME, TO S.AUREUS CAPB AND S.AGALACTIAE CPSC.

```

```

CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: AE000200; AAC74066.1; -.
CC CC DR EMBL: D90735; BAA35746.1; -.
CC CC DR EMBL: AJ238695; CAB43868.1; -.
CC CC DR EMBL: M58708; -. NOT_ANNOTATED_CDS.
CC CC KM Ecocore: Bg1826; etc.
CC CC FT TRANSMEM 33 53
CC CC FT TRANSMEM 92 92 Q -> L (IN STRAIN E2348/69).
CC CC FT VARIANT 169 169 E -> G (IN STRAIN E2348/69).
CC CC FT VARIANT 216 216 E -> G (IN STRAIN E2348/69).
CC CC FT VARIANT 726 726 E -> EEN (IN REF. 2).
CC CC SQ SEQUENCE 726 AA: 81241 MW: BAA060F59680DA22 CRC64;

Query Match
Best Local Similarity 61.3%; Score 38; DB 1; Length 726;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVO 13
Db 243 TRILNSIANNVQ 255

RESULT 8
MISC_DICDI STANDARD: PRT: 1181 AA.
ID MYSC_DICDI
AC P42522;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MYOSIN IC HEAVY CHAIN.
GN MYOC OR DMIC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RC MEDLINE=95348228; PubMed=7622596;
RA Peterson M.D., Novak K.D., Reedy M.C., Ruman J.L., Titus M.A.;
RT "Molecular genetic analysis of myoc, a Dictyostelium myosin I.";
RL J. Cell Sci. 108:1093-1103(1995).
CC CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC CC -1- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY
CC CC AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
CC CC -1- SIMILARITY: BELONGS TO THE MYOSIN TYPE I FAMILY (SMALL MYOSINS).
CC CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: L35323; AAC37427.1; -.
CC CC DR HSSP: P08799; IMND.
CC CC DR Dictydb: DD01090; myoc.
CC CC DR InterPro: IPR001452; -.
CC CC DR InterPro: IPR001609; -.
CC CC DR Pfam: PF00018; SH3; 1.
CC CC DR Pfam: PF00063; myosin_head; 2.

```

DR PRINTS: PRO0193; MYOSINHEAVY.
 DR PRINTS: PRO0452; SH3DOMAIN.
 DR PROSITE: PS00002; SH3; 1.
 KM Myosin: Actin-binding; Arp-binding; Chemotaxis; SH3 domain;
 KM Multigene family.
 FT DOMAIN 1 ? GLOBULAR HEAD-LIKE DOMAIN.
 FT DOMAIN 2 ? NON ALPHA-HELICAL, C-TERMINAL DOMAIN.
 FT NP_BIND 109 116 ARP (POTENTIAL).
 FT DOMAIN 1122 1181 SH3.
 SO SEQUENCE 1181 AA; 132915 MW; 5B1EB47F0CA8803 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 1181;
 Best Local Similarity 53.8%; Pred. No. 35;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYVQ 13
 Db 1167 TKOIGMLPSNYVQ 1179

RESULT 9

PNXL_XENLA STANDARD; PRT; 416 AA.
 AC P49263;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PENTRAXIN FUSION PROTEIN PRECURSOR.
 GN PNXL.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94052237; PubMed=7694301;
 RA Seery L.T., Schoenberg D.R., Barbaux S., Sharp P.M., Whitehead A.S.;
 RT "Identification of a novel member of the pentraxin family in Xenopus
 laevis."
 RT Proc. R. Soc. Lond., B, Biol. Sci. 253:263-270(1993).
 RL -1- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L19881; AAC38013.1; -
 DR HSSP; P02743; IIGN.
 DR InterPro: IPR001759; -
 DR Pfam: PF00354; pentaxin; 1.
 DR PRINTS: PRO0895; PENTAXIN.
 DR PROSITE: PS00289; PENTAXIN; 1.
 KM Pentaxin; Signal.
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 15 416 PENTRAXIN FUSION PROTEIN.
 FT DOMAIN 232 416 PENTAXIN.
 FT DISULFID 251 311 BY SIMILARITY.
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 416 AA; 47241 MW; 62ED645E913A2BD7 CRC64;

Query Match 58.1%; Score 36; DB 1; Length 416;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 GSIASNYVQ 13
 Db 64 GSLSAANYME 72

RESULT 10

CET1_YEAST STANDARD; PRT; 549 AA.
 ID CET1_YEAST
 AC 013297; Q12197;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MRNA CAPPING ENZYME BETA SUBUNIT (POLYNUCLEOTIDE 5'-TRIPHOSPHATASE)
 DE (EC 3.1.3.33) (MRNA 5'-TRIPHOSPHATASE) (TPASE).
 GN CET1 OR YPL228W OR P1433.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180;
 RX MEDLINE=98005091; PubMed=9345280;
 RA Tsukamoto T., Shibagaki Y., Imajobi-Ohtsuki S., Murakoshi T., Suzuki M.,
 RA Nakamura A., Gotoh H., Mizumoto K.;
 RT "Isolation and characterization of the yeast mRNA capping enzyme beta
 RT subunit gene encoding RNA 5'-triphosphatase, which is essential for
 RT cell viability."
 RL Biochem. Biophys. Res. Commun. 239:116-122(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ureastarzu L.A., Vissers S.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=85054827; PubMed=6389537;
 RA Itoh N., Mizumoto K., Kaziro Y.;
 RT "Messenger RNA guanylyltransferase from Saccharomyces cerevisiae. I.
 RT Purification and subunit structure."
 RL J. Biol. Chem. 259:13923-13929(1984).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=85054828; PubMed=6094533;
 RA Itoh N., Mizumoto K., Kaziro Y.;
 RT "Messenger RNA guanylyltransferase from Saccharomyces cerevisiae. II.
 RT Catalytic properties."
 RL J. Biol. Chem. 259:13930-13936(1984).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=87137414; PubMed=3029058;
 RA Itoh N., Yamada H., Kaziro Y., Mizumoto K.;
 RT "Messenger RNA guanylyltransferase from Saccharomyces cerevisiae.
 RT Large scale purification, subunit functions, and subcellular
 RT localization."
 RL J. Biol. Chem. 262:1989-1995(1987).
 RL -1- FUNCTION: FIRST STEP OF M-RNA CAPPING. CONVERTS THE 5'-
 CC TRIPHOSPHATE END OF A NASCENT MRNA CHAIN INTO A DIPHOSPHATE END.
 CC -1- CATALYTIC ACTIVITY: 5'-PHOSPHOPOLYNUCLEOTIDE + H(2)O =
 CC POLYNUCLEOTIDE + ORTHOPHOSPHATE.
 CC -1- COFACTOR: REQUIRES DIVALENT IONS.
 CC -1- SUBUNIT: THE M-RNA CAPPING ENZYME IS COMPOSED OF TWO SEPARATE
 CC CHAINS ALPHA AND BETA, RESPECTIVELY A MRNA GUANYLYLTRANSFERASE AND
 CC AN RNA 5'-TRIPHOSPHATASE.
 CC -1- SIMILARITY: BELONGS TO THE FUNGAL TPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL; AB008799; BAA23522.1; -
DR EMBL; Z73584; CAA97944.1; -
DR EMBL; Z73583; CAA97943.1; -
DR EMBL; X94561; CAA64259.1; -
DR SCD; S0006149; CEM1
KM Hydrolyase; mRNA processing; mRNA capping; Nuclear protein.
FT CONFLICT 242 242 K -> R (IN REF. 2).
SQ SEQUENCE 549 AA; 61821 MW; DBA462FFB7E027F9 CRC64;

Query Match
Best Local Similarity 58.1%; Score 36; DB 1; Length 549;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SSCGSIASNVY 12
Db 162 MEGSIASNVY 171

RESULT 11
HCK_RAT
ID HCK_RAT STANDARD; PRT; 503 AA.
AC P50545; Q64647;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TYROSINE-PROTEIN KINASE HCK (EC 2.7.1.112) (P56-HCK) (HEMOPOIETIC CELL KINASE).
GN HCK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92109719; PubMed=1764064;
RA Okano Y., Sugimoto Y., Fukuoaka M., Matsui A., Nagata K.I., Nozawa Y.;
RT "Identification of rat cDNA encoding hck tyrosine kinase from megakaryocytes.";
RL Blochem. Biophys. Res. Commun. 181:1137-1144(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR; TISSUE=Spleen;
RA Vijaya Gourli B.S., Renu V., Kamatkar S., Swarup G.;
RT "Nucleotide sequence of a cDNA coding for rat hck tyrosine kinase and characterization of its gene product.";
RL J. Biosci. 19:117-129(1994).
CC -1- FUNCTION: MAY SERVE AS PART OF A SIGNALING PATHWAY COUPLING THE FC RECEPTOR TO THE ACTIVATION OF THE RESPIRATORY BURST. MAY ALSO CONTRIBUTE TO NEUTROPHIL MIGRATION AND MAY REGULATE THE DEGRADATION PROCESS OF NEUTROPHILS.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP + PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH MEMBRANES.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC DOMAIN. BELONGS TO THE SRC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S7141; AAB20754.1; -
DR EMBL; M83666; AAA41312.1; -
DR EMBL; X62345; CAA44218.1; -
DR HSSP; P08631; 2HCK.
DR InterPro; IPR000719; -

```

```

DR InterPro; IPR000980; -
DR InterPro; IPR001245; -
DR InterPro; IPR001452; -
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PRO0109; TYRKINASE.
DR PRINTS; PRO0401; SH2DOMAIN.
DR PRINTS; PRO0452; SH3DOMAIN.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
KM Transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;
KW Myristate; SH2 domain; SH3 domain; Alternative initiation.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
FT DOMAIN 55 115 SH3.
FT DOMAIN 121 218 SH2.
FT DOMAIN 239 492 PROTEIN KINASE.
FT NP_BIND 245 253 ATP (BY SIMILARITY).
FT BINDING 267 267 ATP (BY SIMILARITY).
FT ACT_SITE 358 358 BY SIMILARITY.
FT MOD_RES 388 388 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 51 51 F -> V (IN REF. 2).
FT CONFLICT 205 205 K -> R (IN REF. 2).
FT CONFLICT 306 306 I -> T (IN REF. 2).
SQ SEQUENCE 503 AA; 57016 MW; A1FC1F3F0EB2FE73 CRC64;

```

```

Query Match
Best Local Similarity 56.5%; Score 35; DB 1; Length 503;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 TRSSGSIASNVY 12
Db 99 TKEGCIPIPSNVY 110

```

```

RESULT 12
ID LYN_HUMAN STANDARD; PRT; 511 AA.
AC P07948;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TYROSINE-PROTEIN KINASE LYN (EC 2.7.1.112).
GN LYN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87172710; PubMed=3561390;
RA Yamaneishi Y., Fukushige S.-I., Samba K., Sukegawa J., Miyajima N.,
RA Matsubara K.-I., Yamamoto T., Toyoshima K.;
RT "The yes-related cellular gene lyn encodes a possible tyrosine kinase similar to p56lck.";
RL Mol. Cell. Biol. 7:237-243(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94171041; PubMed=8125304;
RA Rider L.G., Raben N., Miller L., Jelsema C.;
RT "The cdms encoding two forms of the lyn protein tyrosine kinase are expressed in rat mast cells and human myeloid cells.";
RL Gene 138:219-222(1994).
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP + PROTEIN TYROSINE PHOSPHATE.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: LYN A (SHOWN HERE) AND LYN B;
CC -1- ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

```

```

CC -I-'SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN, BELONGS TO THE SRC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC DR EMBL; M16038; AAA59540.1; -
CC DR EMBL; M79321; AAB50019.1; -
CC DR PIR; A26719; TVHULY.
CC DR HSSP; P08631; 2HCK.
CC MIM: 165120; -
CC
CC DR InterPro; IPR000719; -
CC DR InterPro; IPR000980; -
CC DR InterPro; IPR001245; -
CC DR InterPro; IPR001452; -
CC DR Pfam; PF00017; SH2; 1.
CC DR Pfam; PF00018; SH3; 1.
CC DR Pfam; PF00069; pkinase; 1.
CC DR PRINTS; PR00109; TYRKINASE.
CC DR PRINTS; PR00452; SH3DOMAIN.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC DR PROSITE; PS50001; SH2; 1.
CC DR PROSITE; PS50002; SH3; 1.
CC
CC K0 Prolo-oncogene; tyrosine-protein kinase; Phosphorylation;
CC KW Transferrase; ATP-binding; Myristate; SH2 domain; SH3 domain;
CC KW Palmitate; Lipoprotein; Alternative splicing.
CC KW INT_MET 0 0
CC FT LIPID 1 1
CC FT LIPID 2 2
CC FT DOMAIN 62 122
CC FT DOMAIN 128 225
CC FT DOMAIN 246 500
CC FT NP_BIND 252 260
CC FT BINDING 274 274
CC FT ACT_SITE 366 366
CC FT MOD_RES 396 396
CC FT MOD_RES 507 507
CC FT VAMPPLIC 22 42
CC SQ SEQUENCE 511 AA; 58442 MW; 8419CD461204E364 CRC64;
CC
CC QY 1 TRSSGSINASYN 12
CC 1: I I I I I I
CC
CC DDB 106 TKKEGPIPSNY 117
CC
CC
CC RESULT 13
CC HCK_MOUSE
CC ID HCK_MOUSE STANDARD; PRT; 524 AA.
CC AC P08103;
CC DT 01-AUG-1988 (Rel. 08, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 30-MAY-2000 (Rel. 35, Last annotation update)
CC DE TYROSINE-PROTEIN KINASE HCK (EC 2.7.1.112) (P56-HCK AND P60-HCK)
CC DE (HEMOPOIETIC CELL KINASE) (B-CELL/MYELOID KINASE) (BMK).
CC HCK.
CC GN Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxId=10090;
CC RN [1]
CC SEQUENCE OF 22-524 FROM N.A.

```

```

RX STAIN-ICR: TISSUE-Macrophage:
RX MEDLINE-88067781: PubMed-3684607:
RT Klemz M.J., Mckercher S.R., Makl R.A.:
RL "Nucleotide sequence of the mouse hck gene."
RN Nucleic Acids Res. 15:9600-9600(1987).
RN [2]
RP SEQUENCE OF 22-524 FROM N.A.
RX MEDLINE-88068587: PubMed-3317404:
RA Holtzman D.A., Cook W.D., Dunn A.R.:
RT "Isolation and sequence of a cDNA corresponding to a src-related gene
RT expressed in murine hemopoietic cells."
RL Proc. Natl. Acad. Sci. U.S.A. 84:8325-8329(1987).
RN [3]
RP SEQUENCE OF 1-22 FROM N.A., AND ALTERNATIVE INITIATION.
RX MEDLINE-91342636: PubMed-1875927:
RA Lock P., Ralph S., Stanley E., Boulet I., Ramsay R., Dunn A.R.:
RT "Two isoforms of murine hck, generated by utilization of alternative
RT translaional initiation codons, exhibit different patterns of
RT subcellular localization."
RL Mol. Cell. Biol. 11:4363-4370(1991).
CC -1- FUNCTION: MAY SERVE AS PART OF A SIGNALING PATHWAY COUPLING THE PC
CC RECEPTOR TO THE ACTIVATION OF THE RESPIRATORY BURST. MAY ALSO
CC CONTRIBUTE TO NEUTROPHIL MIGRATION AND MAY REGULATE THE
CC DEGRADATION PROCESS OF NEUTROPHILS.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: P59-HCK AND P56-HCK ARE ASSOCIATED WITH
CC MEMBRANES. P59-HCK IS ALSO CYTOPLASMIC.
CC -1- ALTERNATIVE PRODUCTS: THE P59-HCK AND P56-HCK ARE PRODUCED BY THE
CC USE OF ALTERNATIVE INITIATION SITES.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CELLS OF THE
CC MELOID AND B-LYMPHOID LINEAGES.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN, BELONGS TO THE SRC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00487; AAA68544.1; -.
DR EMBL: J03023; AAA37305.1; -.
DR PIR: A27282; TVMSHC.
DR PIR: A39973; A39973.
DR HSSP: P08631; 2HCK.
DR MGD: MGI:96052; HCK.
DR InterPro: IPR000719; -.
DR InterPro: IPR000980; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR001452; -.
DR Pfam: PF000017; SH2; 1.
DR Pfam: PF000018; SH3; 1.
DR Pfam: PF000093; PKINASE; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00452; SH3DOMAIN.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
KW transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;
KW Myristate; SH2 domain; SH3 domain; Alternative Initiation.
FT CHAIN 1 524
FT INIT_MET 22 22 TYROSINE-PROTEIN KINASE P59-HCK.
FT LIPID 2 2 FOR P56-HCK.
FT LIPID 23 23 MYRISTATE (BY SIMILARITY).
FT LIPID 23 23 MYRISTATE (BY SIMILARITY).

```

```

FT DOMAIN 76 136 SH3.
FT DOMAIN 142 239 SH2.
FT DOMAIN 260 513 PROTEIN KINASE.
FT NP_BIND 266 274 ATP (BY SIMILARITY).
FT BINDING 288 288 ATP (BY SIMILARITY).
FT ACT_SITE 379 379 BY SIMILARITY.
FT MOD_RES 409 409 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 524 AA; 59129 MW; DF2FD69B38C9706 CRC64;

Query Match 56.5%; Score 35; DB 1; Length 524;
Best Local Similarity 58.3%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 TRSSGSIASNVY 12
   1 1 1 1 1 1 1
Db 120 TRKEGYIPSNVY 131

RESULT 14
FVN_XIPHE STANDARD; PRT; 536 AA.
ID FVN_XIPHE
AC P27446;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FVN (EC 2.7.1.112) (P59-FVN).
GN FVN.
OS Xiphophorus helleri.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphia; Acanthopterygii; Perciformes; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_Taxid=8084;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIO LANCETIIA;
RX MEDLINE=91187435; PubMed=1707152;
RA Hamblin G., Ottile S., Scharl M.;
RT "Conservation of structure and expression of the c-yes and fyn genes
   in lower vertebrates.";
RL Oncogene 6:361-369(1991).
CC -1- FUNCTION: IMPLICATED IN THE CONTROL OF CELL GROWTH.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
CC -1- SUBUNIT: ASSOCIATES THROUGH ITS SH3 DOMAIN, TO THE P85 SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN, BELONGS TO THE SRC SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X54971; CAA38715.1; -
DR HSSP; P06241; IAOV.
DR InterPro; IPR000719; -
DR InterPro; IPR000980; -
DR InterPro; IPR001245; -
DR InterPro; IPR001452; -
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; Kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

```

```

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
KW Proto-oncogene; Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; Myristate; SH3 domain; SH2 domain; Palmitate;
KW Lipoprotein.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT LIPID 2 2 PALMITATE (BY SIMILARITY).
FT LIPID 5 5 PALMITATE (BY SIMILARITY).
FT DOMAIN 81 142 SH3.
FT DOMAIN 148 245 SH2.
FT DOMAIN 270 523 PROTEIN KINASE.
FT MOD_RES 11 11 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT NP_BIND 276 284 ATP (BY SIMILARITY).
FT BINDING 298 298 ATP (BY SIMILARITY).
FT ACT_SITE 389 389 BY SIMILARITY.
FT MOD_RES 419 419 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 530 530 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 536 AA; 60316 MW; 77986A746C3876B CRC64;

Query Match 56.5%; Score 35; DB 1; Length 536;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TRSSGSIASNVY 12
   1 1 1 1 1 1 1
Db 126 TCGSGYIPSNVY 137

RESULT 15
YES_MOUSE STANDARD; PRT; 541 AA.
ID YES_MOUSE
AC Q04736;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE YES (EC 2.7.1.112) (P61-YES)
DE (C-YES).
GN YES.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91173515; PubMed=8437854;
RA Klages S., Adam D., Elseman E., Fargnoli J., Dymecki S.M.,
RA Desiderio S.V., Bolen J.B.;
RT "Molecular cloning and analysis of cDNA encoding the murine c-yes
   tyrosine protein kinase.";
RL Oncogene 8:713-719(1993).
[2]
RP SEQUENCE OF 391-458 FROM N.A.
RX MEDLINE=94266162; PubMed=8206383;
RA Hebert B., Bergeron J., Tiljssen P., Potworowski E.F.;
RT "Protein tyrosine kinases transcribed in a murine thymic medullary
   epithelial cell line.";
RL Gene 143:257-260(1994).
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN, BELONGS TO THE SRC SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: X67677; CAA47909.1; -;
 DR EMBL: I25762; AAA40020.1; -;
 DR PIR: S31645; S31645.
 DR HSSP: P00523; 2PTR.
 DR MGD: MG1:99147; Yes.
 DR InterPro: IPR000719; -;
 DR InterPro: IPR000980; -;
 DR InterPro: IPR001245; -;
 DR InterPro: IPR001452; -;
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00069; PKinase; 1.
 DR PRINTS: PRO0109; TYRKINASE.
 DR PRINTS: PRO0401; SH2DOMAIN.
 DR PRINTS: PRO0452; SH3DOMAIN.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 1.
 KM Proto-oncogene; Tyrosine-protein kinase; Phosphorylation;
 Transferrase; ATP-binding; Myristate; SH3 domain; SH2 domain.
 Lipid 2
 FT DOMAIN 89 150 SH3.
 FT DOMAIN 156 253 SH2.
 FT DOMAIN 275 528 PROTEIN KINASE.
 FT NP_BIND 281 289 ATP (BY SIMILARITY).
 FT BINDING 303 303 ATP (BY SIMILARITY).
 FT ACT_SITE 394 394 BY SIMILARITY.
 FT MOD_RES 424 424 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 372 372 V -> I (IN REF. 2).
 FT CONFLICT 374 390 MISSING (IN REF. 2).
 SO SEQUENCE 541 AA: 60630 MM: 9A773C39D2119EA6 CRC64;

Query Match 56.5%; Score 35; DB 1; Length 541;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TRSGSISANV 12
 1 11 1 1111
 Db 134 TGKSGYIPSNV 145

Search completed: November 19, 2001, 08:25:35
 Job time: 723 sec

THIS PAGE BLANK (USPTO)

Query Match	Best Local Similarity	Score 42: DB 5;	Length 517;
Matches 9;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;
1 TRSSGSIASNY 12			
108 TRSEKIPSNV 119			
RESULT 2			
09V9J3	PRELIMINARY:	PRY:	517 AA.
09V9J3:			
01-MAY-2000 (TREMBlrel. 13, Created)			
01-MAY-2000 (TREMBlrel. 13, Last sequence update)			
01-MAR-2001 (TREMBlrel. 16, Last annotation update)			
SRCA2A PROTEIN.			
GN			
SRCA2A OR CG7873.			
Drosophila melanogaster (Fruit fly).			
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
Ephydroidea; Drosophilidae; Drosophila.			
NCBI_TaxID=7227;			
SEQUENCE FROM N.A. PUB.			
STRAIN=BERKELEY;			
MEDLINE=20196006; PubMed=10731132;			
Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
Brandon R.C., Rogers J., Blazey R.G., Champe M., Pfeiffer B.D.,			
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,			
Borkova D., Botchan M.R., Bouck B.P., Brestein P., Brothier P.,			
Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
Cherry J.M., Chanley S., Dahlke C., Davenport L.B., Davies P.,			
de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
Doudon K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
Durbin K.J., Evangelista C.C., Ferriz C., Fertler S., Fleischmann W.,			
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
Glocke A., Gong E., Gorrell J.H., Gu Z., Guan P., Harris M.,			
Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ideyama C.,			
Jatall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
Liu X., Matcel B., McIntosh T.C., McLeod M.P., McPherson D.,			
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshneft A.,			
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,			
Palazzoio M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,			
Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,			
Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
Syrkasas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,			
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,			
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;			
"The genome sequence of <i>Drosophila melanogaster</i> ."			
Science 287:2195-2195(2000).			
EMBL: AE003784; AA57295.1; -.			
DR	HSSP; P00523; ISRL.		
DR	FLYBase; FBgn0004603; SRC42A.		
DR	InterPro; IPR000719; -.		
DR	InterPro; IPR000980; -.		

[illegible]

RESULT	3		PRT:	778 AA.
ID	014244	PRELIMINARY:		
AC	014244:			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DR	01-JUN-1998 (TREMBLrel. 06, last sequence update)			
DE	01-MAY-2000 (TREMBLrel. 13, last annotation update)			
GN	HYPOTHETICAL 86.1 KDA PROTEIN C6F6.13C IN CHROMOSOME I.			
OS	SCHIZOSACCHAROMYCES POMBE (Pission yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
CC	Schizosaccharomycetales; Schizosaccharomycetaceae;			
OX	NCBI_TaxID=4896;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=972;			
RA	Gentiles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.,			
RL	Submitted (sep-1997) to the EMBL/Genbank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).			
CC	-1- SIMILARITY: TO YEAST YFL034W AND C.ELEGANS F35D11.3.			
DR	EMBL; Z98981; CAB11736.1; -			
DR	InterPro: IPRO00734; -			
KW	PROSITE: PS00120; LIPASE_SRF; UNKNOWN_1.			
FT	Hypothetical protein; Transmembrane.			
FT	TRANSMEM 357 377 POTENTIAL.			
FT	TRANSMEM 378 398 POTENTIAL.			
FT	TRANSMEM 401 421 POTENTIAL.			
FT	TRANSMEM 470 490 POTENTIAL.			
FT	TRANSMEM 512 532 POTENTIAL.			
FT	TRANSMEM 681 701 POTENTIAL.			
SC	SEQUENCE 778 AA; 86129 MW; 0B0B2AC38E5D1796 CRC64;			
QY	3 SSGSIASNYVO 13			
Dd	I IIII IIII			
Db	39 SKGSITENYVQ 49			
Query Match	64.5%;	Score 40;	DB 3;	Length 778;
Best Local Similarity	72.7%;	Pred. No. 25;		
Matches 8;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
RESULT	4	PRELIMINARY:	PRT:	121 AA.
ID	Q9H4Y2			
Q9H4Y2				


```

ID 09XNR3      PRELIMINARY;      PRT;      606 AA.
AC 09XNR3;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Y51A2D.4 PROTEIN.
GN Y51A2D.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN
RP SEQUENCE FROM N.A.
RA McMurtry A.;
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA MEDLINE=94150718; PubMed=7906398;
RX Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Garner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kerhaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sheldron N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.
RL
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AL021497; CAI6400.1; -
DR InterPro: IPR001066; -
DR Pfam: PF00083; sugar_tr. 1
DR PRINTS: PR00171; SUGTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 606 AA; 66658 MW; 5A217A21AF87B501 CRC64;

```

Query Match 61.3%; Score 38; DB 5; Length 606;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 TRSSGSIASNY 11
DB 442 TRNSGSIANNF 452

```

RESULT 9

```

ID 09V5B0      PRELIMINARY;      PRT;      762 AA.
AC 09V5B0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CG1776 PROTEIN.
GN CG1776.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

```

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers T.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bendale J., Beyers R., Binkley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT *The genome sequence of Drosophila melanogaster.
RL
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AE003832; AAF58906.1; -
DR HSSP: P28523; IAGO.
DR FLYBase: FBgn0033441; CG1776.
DR InterPro: IPR000719; -
DR InterPro: IPR002290; -
DR Pfam: PF00069; pkinase. 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM. 1.
DR PROSITE: PS00106; PROTEIN_KINASE_ST. 1.
DR SMART: SM00220; S_TKc. 1.
KW ATP-binding, Serine/threonine-protein kinase, Transferrase.
SQ SEQUENCE 762 AA; 83776 MW; E447DFEC9AE68ED CRC64;

```

Query Match 61.3%; Score 38; DB 5; Length 762;
Best Local Similarity 58.3%; Pred. No. 60;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 TRSSGSIASNY 12
DB 608 TRSSNTSSSYI 619

```

RESULT 10

```

ID 09H3D8      PRELIMINARY;      PRT;      1153 AA.
AC 09H3D8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MSTRP063.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN
RP SEQUENCE FROM N.A.
RA TISSUE=HEART;

```

RA Zhao B., Xu Y.Y., Liu Y.Q., Wang X.Y., Lui B., Ye J., Song L.,
 RA Zhao Y., Cao H.Q., Zhao X.W., Zhang C.L., Zhang J., Liu L.S.,
 RA Ding J.F., Gao R.L., Wu Q.Y., Qiang B.Q., Yuan J.G., Liew C.C.,
 RA Zhao M.S., Hui R.T.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF119814; AAC47945.1; -
 SQ SEQUENCE 1153 AA; 130800 MW; 327FEA675028DA59 CRC64;

Query Match 61.3%; Score 38; DB 4; Length 1153;
 Best Local Similarity 58.3%; Pred. No. 94;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 RSSGSIASNYVQ 13
 :||||| :|||
 Db 108 KNSGSIKNDNYR 119

RESULT 11
 O9P274 PRELIMINARY; PRT; 1265 AA.
 ID O9P274:
 AC O9P274:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE KIA1454 PROTEIN (FRAGMENT).
 GN KIA1454.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20277482; PubMed=10819331;
 RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human
 RT genes. XVII. The complete sequences of 100 new cDNA clones from brain
 RT which code for large proteins in vitro.";
 RL DNA Res. 7:143-150(2000).
 DR EMBL: AB040887; BAA95978.1; -
 DR InterPro: IPR000822; -
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 1265 AA; 142701 MW; 4359125293B25862 CRC64;

Query Match 61.3%; Score 38; DB 4; Length 1265;

Best Local Similarity 58.3%; Pred. No. 1e+02; Mismatches 2; Indels 0; Gaps 0;

OY 2 RSSGSIASNYVQ 13
 :||||| :|||
 Db 370 KNSGSIKNDNYR 381

RESULT 12

O9X331 PRELIMINARY; PRT; 236 AA.

ID O9X331:
 AC O9X331:
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE PX01-61.
 OS Bacillus anthracis.
 OC Bacillus anthracis.
 OC Plasmid virulence plasmid PX01.
 OC Bacillus/Firmicutes; Bacillus/Clostridium group;
 OC NCBI_TaxID=1392;
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=STERNE;
 RA Okinaka R.T., Cloud K., Hamton O., Hoffmaster A., Hill K.K., Keim P.,
 RA Koehler T., Lamke G., Kumano S., Mahillon J., Manter D., Martinez Y.,

RA Ricke D.O., Svensson R., Jackson P.J.;
 RT "The sequence and organization of pX01, the large Bacillus anthracis
 RT plasmid harboring the Anthrax toxin genes.";
 RL J. Bacteriol. 0:0-0(1999).
 DR EMBL: AF065404; AAD32365.1; -
 KW Plasmid.
 SQ SEQUENCE 236 AA; 27161 MW; BE14C0A755887695 CRC64;

Query Match 59.7%; Score 37; DB 2; Length 236;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 RSSGSIASNYVQ 13
 :||||| :|||
 Db 25 KDSGSLFNYVE 36

RESULT 13
 O9LMO0 PRELIMINARY; PRT; 140 AA.
 ID O9LMO0:
 AC O9LMO0:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE F7H2.18 PROTEIN.
 GN F7H2.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Liu S.X., Sakano H., Yu G., Eguu P., Lee J., Lenz C., Pham P.,
 RA Tortum M., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,
 RA Vaynsberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
 RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
 RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT "The sequence of BAC F7H2 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC034256; AAF82154.1; -
 DR InterPro: IPR001680; -
 DR Pfam: PF00400; WD40; 3.
 SQ SEQUENCE 140 AA; 15056 MW; D2B0C857456D248C CRC64;

Query Match 58.1%; Score 36; DB 10; Length 140;

Best Local Similarity 88.9%; Pred. No. 23; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRSSGSIAS 9
 :||| :|||
 Db 56 TRSDSISAS 64

RESULT 14

O54288 PRELIMINARY; PRT; 294 AA.

ID O54288:
 AC O54288:
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE 2-KETO-3-DEOXY GLUCONATE ALDOLASE.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1616;
 RA Kydd C.L., Danson M.J., Hough D.W., Reeve C.D.;

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ224174; CAA11866.1; -
 DR InterPro: IPR002220; -
 DR Pfam: PF00701; DHPS; 1.
 DR ProDom: PD001859; -; 1.
 SQ SEQUENCE 294 AA; 33108 MW; D709A94ACF68B0EF CRC64;

Search completed: November 19, 2001, 08:25:03
 Job time: 741 sec

Query Match 58.1%; Score 36; DB 1; Length 294;
 Best Local Similarity 58.3%; Pred. No. 52;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSGSIASNYVQ 12
 :|::|||
 Db 236 SRIFGSLSSNYVQ 247

RESULT 15

ID Q9LHE5 PRELIMINARY; PRT: 308 AA.
 AC Q9LHE5;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
 DE MITOCHONDRIAL IMPORT RECEPTOR SUBUNIT TOM40 (TRANSLOCASE OF OUTER
 DE MEMBRANE 32 KDA SUBUNIT) (TOM40 PROTEIN).
 GN MZL19.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20363099; PubMed=10907853;
 RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
 RT TAC and BAC clones.";
 RL DNA Res. 7:217-221(2000).
 RN [2]
 RP SEQUENCE OF 1-15; 21-42; 63-80; 207-219 AND 245-258.
 RA Werhahn W., Memeyer A., Jaensch L., Kruff K., Schmitz U.K.,
 RA Braun H.P.;
 RT "Purification and characterization of the preprotein translocase of
 RT the outer mitochondrial membrane from Arabidopsis thaliana:
 RT Identification of multiple forms of TOM20.";
 RL Plant. Physiol. 0:0-0(2000).
 CC -1- FUNCTION: ESSENTIAL FOR THE IMPORT OF PROTEIN PRECURSORS INTO THE
 CC MITOCHONDRION.
 CC -1- SUBUNIT: FORMS PART OF MITOCHONDRIAL RECEPTOR COMPLEX.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC OUTER MEMBRANE.
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC INTRON/EXON MISASSIGNMENT.
 DR EMBL: AP002050; BAB03165.1; ALT_SEQ.
 KW Transport; Protein transport; Outer membrane; Mitochondrion;
 KW Transmembrane.
 FT INT MET 0 0
 FT DOMAIN 1 70 INTERMEMBRANE (POTENTIAL).
 FT TRANSMEM 71 86 POTENTIAL.
 FT DOMAIN 87 187 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 308 AA; 34119 MW; A013F8798BAANA720 CRC64;

Query Match 58.1%; Score 36; DB 10; Length 308;
 Best Local Similarity 54.5%; Pred. No. 54;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SSGSIASNYVQ 13
 :|::|||
 Db 216 STGAVVMNYVQ 226

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:12:35 ; Search time 77.92 Seconds

(Without alignments)
10.114 Million cell updates/sec

Title: US-09-610-118-64
Perfect score: 62
Sequence: 1 TRSSGSIASNYVQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDSB8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSB8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSB8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSB8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSB8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSB8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSB8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSB8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSB8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSB8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSB8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSB8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSB8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSB8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSB8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSB8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSB8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSB8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSB8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSB8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSB8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSB8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	13	22	AA1980.DAT.*
2	50	80.6	13	18	AA1981.DAT.*
3	50	80.6	109	18	AA1982.DAT.*
4	50	80.6	109	18	AA1983.DAT.*
5	39	62.9	13	13	AA1984.DAT.*
6	39	62.9	13	14	AA1985.DAT.*
7	39	62.9	13	21	AA1986.DAT.*
8	39	62.9	111	21	AA1987.DAT.*
9	39	62.9	128	13	AA1988.DAT.*
10	39	62.9	128	13	AA1989.DAT.*
11	39	62.9	128	13	AA1990.DAT.*

12	39	62.9	130	13	AA1991.DAT.*	Recombinant light
13	39	62.9	130	13	AA1992.DAT.*	Recombinant light
14	39	62.9	130	13	AA1993.DAT.*	Recombinant light
15	39	62.9	130	13	AA1994.DAT.*	Recombinant light
16	39	62.9	130	13	AA1995.DAT.*	Recombinant light
17	39	62.9	130	13	AA1996.DAT.*	Recombinant light
18	39	62.9	130	14	AA1997.DAT.*	Aglycosylated anti
19	39	62.9	132	13	AA1998.DAT.*	Recombinant light
20	39	62.9	132	13	AA1999.DAT.*	Recombinant light
21	39	62.9	132	13	AA2000.DAT.*	Recombinant light
22	39	62.9	134	13	AA2001.DAT.*	Recombinant light
23	39	62.9	134	13	AA2002.DAT.*	Recombinant light
24	39	62.9	134	13	AA2003.DAT.*	Recombinant light
25	39	62.9	136	13	AA2004.DAT.*	Recombinant light
26	39	62.9	136	13	AA2005.DAT.*	Recombinant light
27	39	62.9	136	13	AA2006.DAT.*	Recombinant light
28	39	62.9	140	13	AA2007.DAT.*	Recombinant light
29	39	62.9	216	21	AA2008.DAT.*	Amino acid sequenc
30	37	59.7	238	21	AA2009.DAT.*	Arabidopsis thalia
31	37	59.7	267	21	AA2010.DAT.*	Arabidopsis thalia
32	37	59.7	309	21	AA2011.DAT.*	Arabidopsis thalia
33	36	58.1	267	21	AA2012.DAT.*	Arabidopsis thalia
34	36	58.1	332	21	AA2013.DAT.*	Arabidopsis thalia
35	36	58.1	332	21	AA2014.DAT.*	Arabidopsis thalia
36	36	58.1	503	20	AA2015.DAT.*	Zeatin O-glucosyl
37	35	56.5	431	20	AA2016.DAT.*	Enterococcus faeca
38	35	56.5	474	20	AA2017.DAT.*	Enterococcus faeca
39	35	56.5	487	21	AA2018.DAT.*	Amino acid sequenc
40	35	56.5	487	21	AA2019.DAT.*	A murine TANGO 216
41	35	56.5	487	21	AA2020.DAT.*	A murine TANGO 216
42	35	56.5	488	21	AA2021.DAT.*	A human TANGO 216
43	35	56.5	488	21	AA2022.DAT.*	A human TANGO 216
44	35	56.5	488	21	AA2023.DAT.*	Anti-platelet glyco
45	34	54.8	13	21	AA2024.DAT.*	

ALIGNMENTS

RESULT 1	AA1980.DAT.*
ID	AA1980.DAT.* standard: Peptide: 13 AA.
XX	AA1980.DAT.*
AC	AA1980.DAT.*
XX	AA1980.DAT.*
DT	04-APR-2001 (first entry)
XX	AA1980.DAT.*
DE	Anti-TANGO 268 scFv CDR, SEQ ID NO: 64.
XX	AA1980.DAT.*
KW	Human: antibody; scFv; CDR; complementarily determining region;
KW	TANGO 268; cardiant; cerebroprotective; cytosolic; anticoagulant;
KW	thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;
KW	platelet membrane glycoprotein receptor; bleeding disorder;
KW	blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;
KW	ischemia; cardiovascular disease; immunological disease; liver disorder;
XX	cancer.
OS	Homo sapiens.
XX	AA1980.DAT.*
PN	WO200100810-A1.
PD	04-JAN-2001.
XX	AA1980.DAT.*
PF	30-JUN-2000: 2000WO-US18152.
XX	AA1980.DAT.*
PR	30-JUN-1999: 99US-0345468.
XX	AA1980.DAT.*
PR	06-DEC-1999: 99US-0454824.
XX	AA1980.DAT.*
XX	14-FEB-2000: 2000US-0503387.
XX	AA1980.DAT.*
XX	(MILL-) MILLENNIUM PHARM INC.
PI	Busfield SJ, Villediel J, Jandrot-Perrus M, Valchenker W, Gill DS;
PI	Qlan MD, Kingsbury G;

XX WPI: 2001-080877/09.
 XX
 PT New genes encoding human platelet-expressed collagen receptor.
 PT glycoprotein VI, and its modulators, useful for preventing, treating
 PT and diagnosing hemorrhagic disorders, thrombotic diseases and
 PT immunological disorders.
 XX
 PS Claim 32: Page 102, 227pp: English.
 XX
 CC The present sequence is given in a specification relating to an isolated
 CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
 CC glycoprotein VI (GPVI), also called TANGO 268. The GPVI polynucleotides
 CC and polypeptides and their modulators, e.g. antisense nucleic acids,
 CC ribozymes and antibodies, are useful for preventing, treating and
 CC diagnosing disorders associated with aberrant expression or activity of
 CC GPVI. These disorders include bleeding disorders
 CC (e.g. thrombocytopenia), blood vessel injury, thrombotic disorders
 CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
 CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
 CC ischemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
 CC infarction), immunological diseases (e.g. platelet disorder) and
 CC embryonic liver disorders. Preferably they are used to prevent acture
 CC cardiac ischemia following angioplasty and metastatic cancers,
 CC especially of the colon and liver.
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 62; DB 22; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYVQ 13
 |||||
 Db 1 trsgsiasnyvq 13

RESULT 2
 AAW16590
 ID AAW16590 standard; Peptide; 13 AA.
 XX
 AC AAW16590;
 XX
 DT 30-NOV-1997 (first entry)
 XX
 DE Anti-RSV F glycoprotein antibody RSVF2-5 light chain VL6 CDR1.
 XX
 KW Respiratory syncytial virus; RSV; monoclonal antibody; CDR;
 KW complementarity determining region; pneumonia; bronchiolitis;
 KW diagnosis; therapy; vaccine; RSVF2-5.
 XX
 OS Homo sapiens.
 XX
 PN WO9710846-A1.
 XX
 PD 27-MAR-1997.
 XX
 PF 18-SEP-1996; 96WO-US14937.
 XX
 PR 18-SEP-1995; 95US-0003931.
 XX
 PA (INVR-) INTRACEL CORP.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Chanock RM, Crowe JE, Gilmour PS, Murphy BR, Plinkington GR;
 XX
 DR WPI: 1997-202621/18.
 XX
 PT Composition comprising respiratory syncytial virus antibody - useful
 PT for treatment or prophylaxis of active disease or infection
 XX
 PS Claim 9; Page 54; 71pp: English.

XX This peptide sequence comprises complementarity determining region
 CC 1 (CDR1) of the light chain VL6 region of a novel neutralising
 CC human monoclonal antibody to the respiratory syncytial virus (RSV).
 CC The antibody, designated RSVF2-5 (ATCC 69909), selectively binds to
 CC an RSV F glycoprotein epitope. DNA encoding RSVF2-5 Fd and light
 CC chain (see AAW6556-57) was isolated from a phage library prepared
 CC from the RNA of peripheral blood lymphocytes of an HIV-1 infected
 CC donor. A claimed pharmaceutical preparation comprises a carrier
 CC and an antibody that includes the RSVF2-5 VH3 CDR3 (AAW16586) and
 CC which may also include the VH3 CDR2 (AAW16584) and/or CDR1 (AAW16582)
 CC or the entire Fd region (AAW16580), or is an Fab fragment and includes
 CC the RSVF2-5 VL6 CDR3 (AAW16594), CDR2 (AAW16592), CDR1 or entire
 CC light chain (AAW16588). The preparation alternatively comprises a
 CC carrier and a vector that includes a nucleotide sequence encoding
 CC the antibody. The preparations can be used for the treatment or
 CC prophylaxis of active RSV disease or infection (claimed), and may
 CC also be used for RSV detection. The antibody binds and neutralises
 CC antigenic subgroups A and B of RSV with high efficiency.
 XX

Query Match 80.6%; Score 50; DB 18; Length 13;
 Best Local Similarity 76.9%; Pred. No. 0.0039;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYVQ 13
 ||:| |||||
 Db 1 trsgsiasnyvq 13

RESULT 3
 AAW16588
 ID AAW16588 standard; Protein; 109 AA.
 XX
 AC AAW16588;
 XX
 DT 30-NOV-1997 (first entry)
 XX
 DE Anti-RSV F glycoprotein antibody RSVF2-5 light chain VL6.
 XX
 KW Respiratory syncytial virus; RSV; monoclonal antibody; CDR;
 KW complementarity determining region; pneumonia; bronchiolitis;
 KW diagnosis; therapy; vaccine; RSVF2-5.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Peptide
 FT 1..19
 FT /label= FR1
 FT /note= "framework region 1"
 FT 20..32
 FT /label= CDR1
 FT /note= "complementarity determining region 1
 FT (Claim 9)"
 FT 33..47
 FT /label= FR2
 FT /note= "framework region 2"
 FT 48..54
 FT /label= CDR2
 FT /note= "complementarity determining region 2
 FT (Claim 8)"
 FT 55..88
 FT /label= FR3
 FT /note= "framework region 3"
 FT 89..98
 FT /label= CDR3
 FT /note= "complementarity determining region 3
 FT (Claim 7)"
 FT 99..109
 FT /label= FR4
 FT /note= "framework region 4"

```

XX      MO9710846-A1.
PN      27-MAR-1997.
XX      18-SEP-1996; 96WO-US14937.
PD      18-SEP-1995; 95US-0003931.
XX      (INTR-) INTRACEL CORP.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA      Chanock RM, Crowe JE, Gilmour PS, Murphy BR, Pilkington GR;
PI      MPI; 1997-202621/18.
DR      N-PSDB; AAT6557.
XX      Composition comprising respiratory syncytial virus antibody - useful
PT      for treatment or prophylaxis of active disease or infection
XX      Claim 10; Page 52-53; 71pp; English.
XX      This polypeptide sequence comprises the light chain VL6 region
CC      of a novel neutralising human monoclonal antibody to respiratory
CC      syncytial virus (RSV). The antibody, designated RSVF2-5 (ATCC
CC      65909), selectively binds to an RSV F glycoprotein epitope. CDNA
CC      encoding the RSVF2-5 Fd VH3 and light chain VL6 (see AAT6556-57)
CC      was isolated from a phage library prepared from RNA extracted from
CC      the peripheral blood lymphocytes of an HIV-1 infected donor. A
CC      claimed pharmaceutical preparation comprises a carrier and an
CC      antibody that includes the RSVF2-5 VH3 CDR3 (AAW16586) and which may
CC      also include the VH3 CDR2 (AAW16584) and/or CDR1 or the entire Fd
CC      Fd region (AAW16580), or is an Fab fragment and further includes the
CC      RSVF2-5 VL6 CDR3 (AAW16594), CDR2 (AAW16592), CDR1 (AAW16590) or entire
CC      light chain (AAW16588). The preparation alternatively comprises a
CC      carrier and a vector that includes a nucleotide sequence encoding
CC      the antibody. The preparations can be used for the treatment or
CC      prophylaxis of active RSV disease or infection (claimed), and may
CC      also be used for RSV detection. The antibody binds and neutralises
CC      antigenic subgroups A and B of RSV with high efficiency.
XX      SQ      Sequence 109 AA;
Query Match      80.6%; Score 50; DB 18; Length 109;
Best Local Similarity 76.9%; Pred. No. 0.046;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY      1 TRSSGSIASNYVQ 13
      ||: | | | | | | |
Db      20 traggriasnyvq 32
RESULT 4
AAW14785
ID      AAW14785 standard; Protein; 109 AA.
XX      AAW14785;
AC      20-JUN-1997 (first entry)
XX      Human monoclonal antibody RSVF2-5 light chain VL6.
DE      Respiratory syncytial virus; RSV; monoclonal antibody; Mab;
KM      diagnosis; prophylaxis; immunotherapy; therapy; Fab; Fd.
XX      Homo sapiens.
OS      Key      Location/Qualifiers
FH      Region      1..19
FT      Region      /label= FR1
FT      Region      20..32
FT      Region      /label= CDR1

```

```

FT      /note= "Claim 9"
FT      Region      33..47
FT      /label= FR2
FT      Region      48..54
FT      /label= CDR2
FT      /note= "Claim 8"
FT      Region      55..88
FT      /label= FR3
FT      Region      89..98
FT      /label= CDR3
FT      /note= "Claim 7"
FT      Region      99..109
FT      /label= FR4
XX      MO9711177-A1.
PN      27-MAR-1997.
XX      18-SEP-1996; 96WO-US14944.
PD      18-SEP-1995; 95US-0003931.
XX      (INTR-) INTRACEL CORP.
XX      Gilmour PS, Pilkington GR;
PI      MPI; 1997-202886/18.
DR      N-PSDB; AAT63418.
XX      Monoclonal antibody specific for respiratory syncytial virus - used
PT      for diagnosis and immuno-prophylaxis or immuno-therapy of RSV
XX      disease
XX      Claim 10; Page 49-50; 68pp; English.
XX      The light chain variable region (AAW14785) and heavy chain variable
CC      region (AAW14784) are provided of a novel, fully human monoclonal
CC      antibody (Mab), designated RSVF2-5, which selectively binds to
CC      the F glycoprotein of respiratory syncytial virus (RSV) and which
CC      neutralises RSV in vivo. To obtain RSVF2-5, a packaged phage
CC      library was prep'd. from amplified heavy and light chain variable
CC      region clones derived from the peripheral blood lymphocytes of an
CC      HIV-1 infected donor. The phage library was panned with RSV
CC      proteins and isolated clones were sequenced. Isolated nucleic
CC      acids (AAT63417-18) can be used to produce novel polypeptides, esp.
CC      Fd or Fab fragments useful in the diagnosis of RSV disease, in
CC      methods for detecting the presence of RSV in a sample, and in the
CC      immunoprophylaxis and immunotherapy of RSV disease.
XX      SQ      Sequence 109 AA;
Query Match      80.6%; Score 50; DB 18; Length 109;
Best Local Similarity 76.9%; Pred. No. 0.046;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY      1 TRSSGSIASNYVQ 13
      ||: | | | | | | |
Db      20 traggriasnyvq 32
RESULT 5
AAR23736
ID      AAR23736 standard; Protein; 13 AA.
XX      AAR23736;
AC      27-OCT-1992 (first entry)
XX      Ligand CDR(d).
DE      Complementarity determining region: heavy chain variable domain;
KM      antigen binding site.

```

XX OS Synthetic.
 XX PA WO9206193-A.
 XX PN
 XX PD 16-APR-1992.
 XX PF 04-OCT-1991; 91WO-GB01726.
 XX PR 05-OCT-1990; 90GB-0021679.
 XX PA (GORM/) GORMAN S D.
 XX PI Gorman SD, Routledge EG, Waldmann H;
 XX DR WPI: 1992-150879/18.
 XX DR N-PSDB: AAQ24340.
 XX PT Ligands and antibodies with binding affinity for CD3 antigen -
 XX PT for treatment of immunosuppression e.g. in graft rejection, and
 XX PT cancer, esp. lymphoid malignancies
 XX PS
 XX PS Disclosure: Page 4; 49pp; English.

CC The sequence given is a complementarily determining region (CDR)
 CC designated CDR(d). CDR's are found in the variable domains of light
 CC and heavy chains which form the antigen binding site, and act as
 CC connectors between the four framework regions.
 CC It has been noted that there seem to be no characteristic features
 CC which distinguish human from mouse or rat CDR's and they are
 CC therefore immunologically identical. The CDR described is used in
 CC a ligand with a binding affinity for the human CD3 antigen within
 CC the framework regions of the heavy chain variable domain along with
 CC the CDR's described in AAR23737 and AAR23738.

SO Sequence 13 AA:

Query Match 62.9%; Score 39; DB 13; Length 13;
 Best Local Similarity 66.7%; Pred. No. 0.46;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYV 12
 1 ||||| : ||||
 Db 1 tlssgsienyny 12

RESULT 6

ID AAR50108 standard; peptide; 13 AA.

AC AAR50108;

DT 27-APR-1994 (first entry)

DE CDR (d).

KW Antibody; glycosylation; IgG; binding affinity; CD3; antigen;
 KW CDR; complementarity determining region; heavy chain; light chain;
 KW variable domain; immunosuppressant; T-cell; mitogenesis;
 KW cytokine; graft; rejection; cancer; proliferation.

OS Synthetic.

PN WO9319196-A.

PD 30-SEP-1993.

PF 21-OCT-1992; 92WO-GB01933.

PR 24-MAR-1992; 92GB-0006422.

PA (BOLT/) BOLT S L.

PA (CLAR/) CLARK M R.
 PA (GORM/) GORMAN S D.
 PA (ROUT/) ROUTLEDGE E G.
 PA (WALD/) WALDMANN H.
 PI Bolt SL, Clark MR, Gorman SD, Routledge EG, Waldmann H;
 DR WPI: 1993-320762/40.

PT Aglycosylated IgG antibody against human CD3 antigen - for
 PT preventing graft rejection and treating cancer, and does not
 PT induce T cell proliferation and cytokine prodn.

PS Claim 3; Page 25; 41pp; English.

CC An aglycosylated IgG antibody having a binding affinity for the human
 CC CD3 antigen complex comprises at least one CDR selected from the amino
 CC acid sequences given in AAR50105-R50110. The heavy and light chain
 CC variable domains of an aglycosylated antibody is given in AAR41721-
 CC AAR41722. The antibodies are useful as immunosuppressants. Unlike
 CC glycosylated antibodies, they do not induce T-cell mitogenesis nor
 CC do they cause high level release of cytokines.

SO Sequence 13 AA:

Query Match 62.9%; Score 39; DB 14; Length 13;
 Best Local Similarity 66.7%; Pred. No. 0.46;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYV 12
 1 ||||| : ||||
 Db 1 tlssgsienyny 12

RESULT 7

ID AAY68804 standard; peptide; 13 AA.

AC AAY68804;

DT 16-MAY-2000 (first entry)

DE Amino acid sequence of CDR (d) of an anti-CD3 IgG antibody.

KW Complementarity determining region; CDR; rat; immunoglobulin G; IgG;
 KW CD3 antigen complex; chimeric antibody; immunosuppression;
 KW CD3 antigen-T-cell receptor complex; graft rejection; cancer.

OS Ratfus sp.

PN WO200005268-A1.

PD 03-FEB-2000.

PF 21-JUL-1999; 99WO-GB02380.

PR 21-JUL-1998; 98GB-0015909.

PA (BYGI-) BTG INT LTD.

PI Waldmann H, Frewin M;

DR WPI: 2000-182655/16.

DR N-PSDB: AA260593.

PT New humanized anti-CD3 antibodies, used for treating cancer or for
 PT immunosuppression and preventing graft rejection -
 PS Claim 1; Page 38; 56pp; English.

CC AAY68801-06 represent complementarity determining regions (CDRs) of a
 CC rat immunoglobulin G (IgG) antibody which is specific for the CD3

CC antigen complex. The CDR regions are used to produce chimeric.
 CC human/rodent anti-CD3 antibodies, which have a rodent CD3 light chain
 CC variable region and a human heavy chain variable region. The anti-CD3
 CC antibodies can render T-cells non-functional by antibody blockade of the
 CC CD3 antigen-T-cell receptor (TCR) complex. They can be used for
 CC immunosuppression, particularly for the control of graft rejection.
 CC The antibodies can also enhance or re-direct T-cell responses to
 CC antigens. They can be used in the treatment of cancer.
 XX
 XX Sequence 13 AA;

Query Match 62.9%; Score 39; DB 21; Length 13;
 Best Local Similarity 66.7%; Pred. NO. 0.46;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TRSGSISANVY 12
 | 1111:1:1111
 Db 1 t1ssgn1ennyv 12

RESULT 8

AAV68807 standard; Protein; 111 AA.
 ID AAV68807 standard; Protein; 111 AA.

AC AAV68807;

DT 16-MAY-2000 (first entry)

DE A light chain variable region of a rat anti-CD3 IgG antibody.

KM CD3 antigen complex; chimeric antibody; immunosuppression; light chain;

KW CD3 antigen-T-cell receptor complex; graft rejection; cancer.

XX Rattus sp.

PN W0200005268-A1.

PD 03-FEB-2000.

PF 21-JUL-1999; 99WO-G802380.

PR 21-JUL-1998; 98GB-0015909.

PA (BTGI-) BTG INT LTD.

PI Waldmann H, Frewin M;

DR WPI; 2000-18265/16.

DR N-PSDB; AAZ60596.

PT New humanized anti-CD3 antibodies, used for treating cancer or for

PT immunosuppression and preventing graft rejection

PS Claim 5; Page 40-41; 56pp; English.

CC The present sequence represents the light chain variable region of a rat
 CC immunoglobulin G (IgG) antibody which is specific for the CD3 antigen
 CC complex. The light chain variable regions is used to produce chimeric
 CC human/rodent anti-CD3 antibodies, which have a rodent CD3 light chain
 CC variable region and a human heavy chain variable region. The anti-CD3
 CC antibodies can render T-cells non-functional by antibody blockade of
 CC the CD3 antigen-T-cell receptor (TCR) complex. They can be used for
 CC immunosuppression, particularly for the control of graft rejection.
 CC The antibodies can also enhance or re-direct T-cell responses to
 CC antigens. They can be used in the treatment of cancer.

XX Sequence 111 AA;

Query Match 62.9%; Score 39; DB 21; Length 111;
 Best Local Similarity 66.7%; Pred. NO. 5.5;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TRSGSISANVY 12
 | 1111:1:1111
 Db 23 t1ssgn1ennyv 34

RESULT 9

AAAR23772 standard; Protein; 128 AA.
 ID AAR23772 standard; Protein; 128 AA.

AC AAR23772;

DT 27-OCT-1992 (first entry)

DE Recombinant light chain variable domain (6).

KM Complementarity determining region; light chain variable domain;
 KW antigen binding site; ligand; framework region; cancer; transplant.

XX Synthetic.

FT Key Location/Qualifiers

FT Region 23..35 /label= CDR(d)

FT Region 51..57 /label= CDR(e)

FT Region 92..98 /label= CDR(e)

PN W09206193-A.

PD 16-APR-1992.

PF 04-OCT-1991; 91WO-G801726.

PR 05-OCT-1990; 90GB-0021679.

PA (GORM/) GORMAN S D.

PI Gorman SD, Routledge EG, Waldmann H;

DR WPI; 1992-150879/18.

PT Ligands and antibodies with binding affinity for CD3 antigen -

PT for treatment of immunosuppression e.g. in graft rejection, and

PS cancer, esp. Lymphoid malignancies

PS Claim 7; Page 31; 49pp; English.

CC The sequence given is a recombinant human light chain variable
 CC domain ligand containing the complementarity determining region
 CC (CDR) given in AAR23776 and AAR23737. CDR's are found in the variable
 CC domains of light and heavy chains which form the antigen binding site,
 CC and act as connectors between the four framework regions.
 CC It has been noted that there seem to be no characteristic features
 CC which distinguish human from mouse or rat CDR's and they are
 CC therefore immunologically identical. This ligand has binding affinity
 CC for the human CD3 antigen and due to the lack of immunological
 CC response caused by the synthetic CDR's the ligand can be considered to
 CC be humanised. This ligand can be used to manufacture medicaments
 CC for use in immunosuppression esp. in patients with cancer or transplant
 CC recipients.

XX Sequence 128 AA;

Query Match 62.9%; Score 39; DB 13; Length 128;
 Best Local Similarity 66.7%; Pred. NO. 6.5;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TRSGSISANVY 12
 | 1111:1:1111
 Db 23 t1ssgn1ennyv 34

```

RESULT 10
ID AAR23778 standard; Protein: 128 AA.
XX
AC AAR23778;
XX
AC AAR23778;
XX
DT 27-OCT-1992 (first entry)
XX
DE Recombinant light chain variable domain (11).
XX
KM Complementarity determining region: light chain variable domain;
KW antigen binding site; ligand; framework region; cancer; transplant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 23..29
FT /label= CDR(e)
FT Region 45..57
FT /label= CDR(d)
FT Region 92..98
FT /label= CDR(e)
XX
PN W09206193-A.
XX
PD 16-APR-1992.
XX
PF 04-OCT-1991; 91WO-GB01726.
XX
PR 05-OCT-1990; 90GB-0021679.
XX
PA (GORM/) GORMAN S D.
PI Gorman SD, Routledge EG, Waldmann H;
XX
DR WPI: 1992-150879/18.
XX
PT Ligands and antibodies with binding affinity for CD3 antigen -
PT for treatment of immunosuppression e.g. in graft rejection, and
PT cancer, esp. lymphoid malignancies
XX
PS Claim 7; Page 31; 49pp; English.
XX
CC The sequence given is a recombinant human light chain variable
CC domain ligand containing the complementarity determining region
CC (CDR) given in AAR23736 and AAR23737. CDR's are found in the
CC variable domains of light and heavy chains which form the antigen
CC binding site, and act as connectors between the four framework regions.
CC It has been noted that there seem to be no characteristic features
CC which distinguish human from mouse or rat CDR's and they are
CC therefore immunologically identical. This ligand has binding affinity
CC for the human CD3 antigen and due to the lack of immunological
CC response caused by the synthetic CDR's the ligand can be considered to
CC be humanised. This ligand can be are used to manufacture medicaments
CC for use in immunosuppression esp. in patients with cancer or transplant
CC recipients.
XX
SQ Sequence 128 AA;

```

Query Match 62.9%; Score 39; DB 13; Length 128;
 Best Local Similarity 66.7%; Pred. No. 6.5;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 TRSSGSIASNYV 12
| |||:| :|||
Db 45 tssgnlennyv 56

```

RESULT 11
 AAR23814

```

ID AAR23814 standard; Protein: 128 AA.
XX
AC AAR23814;
XX
AC AAR23814;
XX
DT 27-OCT-1992 (first entry)
XX
DE Recombinant light chain variable domain (13).
XX
KM Complementarity determining region: light chain variable domain;
KW antigen binding site; ligand; framework region; cancer; transplant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 23..29
FT /label= CDR(e)
FT Region 45..51
FT /label= CDR(e)
FT Region 86..98
FT /label= CDR(d)
XX
PN W09206193-A.
XX
PD 16-APR-1992.
XX
PF 04-OCT-1991; 91WO-GB01726.
XX
PR 05-OCT-1990; 90GB-0021679.
XX
PA (GORM/) GORMAN S D.
PI Gorman SD, Routledge EG, Waldmann H;
XX
DR WPI: 1992-150879/18.
XX
PT Ligands and antibodies with binding affinity for CD3 antigen -
PT for treatment of immunosuppression e.g. in graft rejection, and
PT cancer, esp. lymphoid malignancies
XX
PS Claim 7; Page 31; 49pp; English.
XX
CC The sequence given is a recombinant human light chain variable
CC domain ligand containing the complementarity determining region
CC (CDR) given in AAR23736 and AAR23737. CDR's are found in the
CC variable domains of light and heavy chains which form the antigen
CC binding site, and act as connectors between the four framework regions.
CC It has been noted that there seem to be no characteristic features
CC which distinguish human from mouse or rat CDR's and they are
CC therefore immunologically identical. This ligand has binding affinity
CC for the human CD3 antigen and due to the lack of immunological
CC response caused by the synthetic CDR's the ligand can be considered to
CC be humanised. This ligand can be are used to manufacture medicaments
CC for use in immunosuppression esp. in patients with cancer or transplant
CC recipients.
XX
SQ Sequence 128 AA;

```

Query Match 62.9%; Score 39; DB 13; Length 128;
 Best Local Similarity 66.7%; Pred. No. 6.5;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 TRSSGSIASNYV 12
| |||:| :|||
Db 86 tssgnlennyv 97

```

RESULT 12
 AAR23740
 ID AAR23740 standard; Protein: 130 AA.
 AC AAR23740;
 AC AAR23740;

```

DT 27-OCT-1992 (first entry)
XX
DE Recombinant light chain variable domain.
XX
KW Complementarity determining region; light chain variable domain;
XX antigen binding site; ligand; framework region; cancer; transplant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 23..35
FT /label= CDR(d)
FT Region 51..57
FT /label= CDR(e)
FT Region 92..100
FT /label= CDR(f)
XX
PN WO9206193-A.
XX
PD 16-APR-1992.
XX
PE 04-OCT-1991; 91WO-GB01726.
XX
PR 05-OCT-1990; 90GB-0021679.
XX
PA (GORM/) GORMAN S D.
XX
PI Gorman SD, Routledge EG, Waldmann H;
XX
DR WPI: 1992-150879/18.
XX
DR N-PSDB; AA024515.
XX
PT Ligands and antibodies with binding affinity for CD3 antigen -
PT for treatment of immunosuppression e.g. in graft rejection, and
XX cancer, esp. lymphoid malignancies
XX
PS Claim 9: Page 31: 49pp; English.
XX
XX
CC The sequence given is a recombinant human heavy chain variable
CC domain ligand containing the complementarity determining regions
CC (CDR's) given in AAR23736, AAR23737 and AAR23738. CDR's are found in
CC the variable domains of light and heavy chains which form the antigen
CC binding site, and act as connectors between the four framework regions.
CC It has been noted that there seem to be no characteristic features
CC which distinguish human from mouse or rat CDR's and they are
CC therefore immunologically identical. This ligand has binding affinity
CC for the human CD3 antigen and due to the lack of immunological
CC response caused by the synthetic CDR's the ligand can be considered to
CC be humanised. This ligand and the one described in AAR23739 are used to
CC manufacture medicaments for use in immunosuppression esp. in patients
CC with cancer or transplant recipients.
XX
SQ Sequence 130 AA:
XX
XX
Query Match 62.9%; Score 39; DB 13; Length 130;
Best Local Similarity 66.7%; Pred. No. 6.6;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 TRSSGSIASNVY 12
DB 23 t1ssgnlennvy 34
XX
RESULT 13
AAR23775
ID AAR23775 standard; Protein; 130 AA.
XX
AC AAR23775;
XX
DE 27-OCT-1992 (first entry)
XX
KW Recombinant light chain variable domain (8).
XX

```

```

XX
KW Complementarity determining region; light chain variable domain;
XX antigen binding site; ligand; framework region; cancer; transplant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 23..35
FT /label= CDR(d)
FT Region 51..59
FT /label= CDR(f)
FT Region 93..100
FT /label= CDR(e)
XX
PN WO9206193-A.
XX
PD 16-APR-1992.
XX
PE 04-OCT-1991; 91WO-GB01726.
XX
PR 05-OCT-1990; 90GB-0021679.
XX
PA (GORM/) GORMAN S D.
XX
PI Gorman SD, Routledge EG, Waldmann H;
XX
DR WPI: 1992-150879/18.
XX
PT Ligands and antibodies with binding affinity for CD3 antigen -
PT for treatment of immunosuppression e.g. in graft rejection, and
XX cancer, esp. lymphoid malignancies
XX
PS Claim 7: Page 31: 49pp; English.
XX
XX
CC The sequence given is a recombinant human light chain variable
CC domain ligand containing the complementarity determining region
CC (CDR) given in AAR23736, AAR23737 and AAR23738. CDR's are found in the
CC variable domains of light and heavy chains which form the antigen
CC binding site, and act as connectors between the four framework regions.
CC It has been noted that there seem to be no characteristic features
CC which distinguish human from mouse or rat CDR's and they are
CC therefore immunologically identical. This ligand has binding affinity
CC for the human CD3 antigen and due to the lack of immunological
CC response caused by the synthetic CDR's the ligand can be considered to
CC be humanised. This ligand can be are used to manufacture medicaments
CC for use in immunosuppression esp. in patients with cancer or transplant
XX recipients.
XX
SQ Sequence 130 AA:
XX
XX
Query Match 62.9%; Score 39; DB 13; Length 130;
Best Local Similarity 66.7%; Pred. No. 6.6;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 TRSSGSIASNVY 12
DB 23 t1ssgnlennvy 34
XX
RESULT 14
AAR23779
ID AAR23779 standard; Protein; 130 AA.
XX
AC AAR23779;
XX
DE 27-OCT-1992 (first entry)
XX
KW Recombinant light chain variable domain (12).
XX
KW Complementarity determining region; light chain variable domain;
XX antigen binding site; ligand; framework region; cancer; transplant.
XX

```

OS	Synthetic.
XX	
FH	Key
FT	Region
FT	Region
FT	Region
FT	Region
FT	Region
FT	Region
FT	Region
PN	MO9206193-A.
PD	16-APR-1992.
PP	04-OCT-1991; 91WO-GB01726.
PR	05-OCT-1990; 90GB-0021679.
PA	(GORM/) GORMAN S D.
PX	Gorman SD, Routledge EG, Waldmann H;
DR	WPI; 1992-150879/18.
PT	Ligands and antibodies with binding affinity for CD3 antigen -
PT	for treatment of immunosuppression e.g. In graft rejection, and
PT	cancer, esp. lymphoid malignancies
PS	Claim 7; Page 31; 48pp; English.
XX	
CC	The sequence given is a recombinant human light chain variable
CC	domain ligand containing the complementarity determining region
CC	(CDM) given in AAR23736, AAR23737 and AAR23738. CDR's are found in the
CC	variable domains of light and heavy chains which form the antigen
CC	binding site, and act as connectors between the four framework regions.
CC	It has been noted that there seem to be no characteristic features
CC	which distinguish human from mouse or rat CDR's and they are
CC	therefore immunologically identical. This ligand has binding affinity
CC	for the human CD3 antigen and due to the lack of immunological
CC	response caused by the synthetic CDR's the ligand can be considered to
CC	be humanised. This ligand can be used to manufacture medicaments
CC	for use in immunosuppression esp. in patients with cancer or transplant
CC	recipients.
SO	Sequence 130 AA;
QY	Query Match 62.9%; Score 39; DB 13; Length 130; Best Local Similarity 66.7%; Pred. No. 6.6; Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
DB	1 TRSSGSIASNYV 12 1 : : 45 tssgnienyy 56
RESULT 15	
AAR23817	
ID	AAR23817 standard; Protein: 130 AA.
AC	AAR23817;
DT	27-OCT-1992 (first entry)
DE	Recombinant light chain variable domain (16).
KW	Complementarity determining region: light chain variable domain;
XX	antigen binding site; ligand; framework region; cancer; transplant.
OS	Synthetic.
XX	
FH	Key
FT	Region
FT	Location/Qualifiers 23..29 /label= CDR(e) 45..57 /label= CDR(d) 92..100 /label= CDR(f)

FT		/label= CDR(e)
FT	Region	45..53
FT		/label= CDR(f)
FT	Region	88..100
FT		/label= CDR(d)
XX		
PN	WO9206193-A.	
XX		
PD	16-APR-1992.	
XX		
PF	04-OCT-1991;	91WO-GB01726.
XX		
PR	05-OCT-1990;	90GB-0021679.
XX		
PA	(GORM/) GORMAN S D.	
XX		
PI	Gorman SD, Routledge EC, Waldmann H;	
XX		
DR	WPI; 1992-150879/18.	
XX		
PT	Ligands and antibodies with binding affinity for CD3 antigen -	
PT	for treatment of immunosuppression e.g. in graft rejection, and	
PT	cancer, esp. lymphoid malignancies	
XX		
PS	Claim 7; Page 31; 49pp; English.	
XX		
CC	The sequence given is a recombinant human light chain variable	
CC	domain ligand containing the complementarity determining region	
CC	(CDR) given in AAR23736, AAR23737 and AAR23738. CDR's are found in the	
CC	variable domains of light and heavy chains which form the antigen	
CC	binding site, and act as connectors between the four framework regions.	
CC	It has been noted that there seem to be no characteristic features	
CC	which distinguish human from mouse or rat CDR's and they are	
CC	therefore immunologically identical. This ligand has binding affinity	
CC	for the human CD3 antigen and due to the lack of immunological	
CC	response caused by the synthetic CDR's the ligand can be considered to	
CC	be humanised. This ligand can be are used to manufacture medicaments	
CC	for use in immunosuppression esp. in patients with cancer or transplant	
CC	recipients.	
XX		
SO	Sequence	130 AA;

Query Match	62.9%;	Score 39;	DB 13;	Length 130;
Best Local Similarity	66.7%;	Pred. No. 6.6;		
Matches	8;	Conservative	2;	Mismatches
			2;	Indels
			0;	Gaps
OY	1 TRSSGSIASNYV 12			0;
	1 : :			
Db	88 tllsgniennyv 99			

Search completed: November 19, 2001, 08:12:35
Job time: 88 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:15 ; Search time 526.89 Seconds
(without alignments)
6.851 Million cell updates/sec

Title: US-09-610-118-64

Perfect score: 62

Sequence: 1 TRSSGSIASNYVQ 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep: *
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep: *
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep: *
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep: *
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep: *
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep: *
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep: *
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep: *
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep: *
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep: *
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep: *
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep: *
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep: *
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep: *
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep: *
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep: *
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep: *
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep: *
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep: *
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep: *
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep: *
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep: *
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep: *
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	13	1	PCT-US00-18152-64
2	62	100.0	13	20	US-09-610-118-64
3	62	100.0	13	22	US-09-832-312-64
4	62	100.0	88	24	US-60-195-053-2358
5	62	100.0	94	24	US-60-195-053-2359
6	62	100.0	100	24	US-60-196-710-4905
7	62	100.0	100	24	US-60-196-718-4971
8	62	100.0	161	1	PCT-US00-265248-6775
9	56	90.3	258	1	PCT-US01-19110-1234

10	56	90.3	258	22	US-09-880-748-1234	Sequence 1234, Ap
11	55	88.7	13	14	US-09-043-522A-11	Sequence 11, Ap1
12	55	88.7	13	14	US-09-043-530-11	Sequence 11, Ap1
13	55	88.7	79	24	US-60-163-062-790	Sequence 790, App
14	55	88.7	79	24	US-60-163-062-948	Sequence 948, App
15	55	88.7	94	24	US-60-163-123-1181	Sequence 1181, App
16	55	88.7	94	24	US-60-163-123-1181	Sequence 1181, App
17	55	88.7	94	24	US-60-234-446-712	Sequence 712, App
18	55	88.7	95	24	US-60-162-243-594	Sequence 594, App
19	55	88.7	95	24	US-60-162-243-594	Sequence 4885, App
20	55	88.7	104	24	US-60-162-247-3052	Sequence 3052, App
21	55	88.7	104	24	US-60-162-247-3052	Sequence 3489, App
22	55	88.7	108	24	US-60-169-840-7362	Sequence 7123, App
23	55	88.7	108	24	US-60-169-840-7362	Sequence 7362, App
24	55	88.7	109	14	US-09-043-522A-9	Sequence 9, Ap11
25	55	88.7	109	14	US-09-043-530-9	Sequence 5171, App
26	55	88.7	109	24	US-60-169-868-5171	Sequence 24, Ap1
27	55	88.7	110	14	US-09-043-522A-24	Sequence 421, App
28	55	88.7	110	14	US-09-043-530-24	Sequence 421, App
29	55	88.7	111	24	US-60-162-243-421	Sequence 4435, App
30	55	88.7	111	24	US-60-162-247-4435	Sequence 8023, App
31	55	88.7	111	24	US-60-169-840-8023	Sequence 5926, App
32	55	88.7	111	24	US-60-169-868-5926	Sequence 909, App
33	53	85.5	253	1	PCT-US01-19110-1125	Sequence 1125, App
34	53	85.5	253	1	PCT-US01-19110-1125	Sequence 1167, App
35	53	85.5	253	1	PCT-US01-19110-1167	Sequence 1375, App
36	53	85.5	253	1	PCT-US01-19110-1375	Sequence 1375, App
37	53	85.5	253	22	US-09-880-748-909	Sequence 1125, App
38	53	85.5	253	22	US-09-880-748-1125	Sequence 1167, App
39	53	85.5	253	22	US-09-880-748-1167	Sequence 1375, App
40	53	85.5	253	22	US-09-880-748-1375	Sequence 1375, App
41	52	83.9	257	22	US-09-880-748-1375	Sequence 1375, App
42	52	83.9	257	22	US-09-880-748-1375	Sequence 1375, App
43	44	71.0	86	21	US-09-760-476-1377	Sequence 824, App
44	44	71.0	86	21	US-09-760-476-824	Sequence 603, App
45	44	71.0	237	21	US-09-760-479-603	

ALIGNMENTS

RESULT 1
PCT-US00-18152-64
Sequence 64, Application PC/TUS0018152
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-211-228
CURRENT APPLICATION NUMBER: PCT/US00/18152
CURRENT FILING DATE: 2000-06-30
EARLIER APPLICATION NUMBER: 09/503,387
EARLIER FILING DATE: 2/14/00
EARLIER APPLICATION NUMBER: 09/454,824
EARLIER FILING DATE: 12/6/99
EARLIER APPLICATION NUMBER: 09/345,468
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 64
LENGTH: 13
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-18152-64

Query Match 100.0%; Score 62; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVQ 13
|||||
DB 1 TRSSGSIASNYVQ 13

```
RESULT 2
US-09-610-118-64
; Sequence 64, Application US/09610118
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Valinchenker, W.
; APPLICANT: Gill, D.
; APPLICANT: Qian, M.
; APPLICANT: Kingsbury, G.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-211
; CURRENT APPLICATION NUMBER: US/09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2/14/00
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 12/6/99
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 6/30/99
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-610-118-64

Query Match          100.0%; Score 62; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
Db 1 TRSSGSIASNYQ 13

RESULT 3
US-09-832-312-64
; Sequence 64, Application US/09832312
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-64

Query Match          100.0%; Score 62; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
Db 1 TRSSGSIASNYQ 13

RESULT 4
US-60-195-053-2358
; Sequence 2358, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2358
; LENGTH: 88
; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-2358

Query Match          100.0%; Score 62; DB 24; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
Db 45 TRSSGSIASNYQ 57

RESULT 5
US-60-195-053-2359
; Sequence 2359, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2359
; LENGTH: 94
; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-2359

Query Match          100.0%; Score 62; DB 24; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
Db 42 TRSSGSIASNYQ 54

RESULT 6
US-60-196-710-4905
; Sequence 4905, Application US/60196710
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: CL000450
; CURRENT APPLICATION NUMBER: US/60/196,710
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7166
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 4905
LENGTH: 100
TYPE: PRT
ORGANISM: HUMAN
US-60-196-710-4905

Query Match 100.0%; Score 62; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
|||||
DB 45 TRSSGSIASNYQ 57

RESULT 7
US-60-196-718-4971
Sequence 4971, Application US/60196718
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
FILE REFERENCE: CL000456
CURRENT APPLICATION NUMBER: US/60/196,718
CURRENT FILING DATE: 2000-04-13
NUMBER OF SEQ ID NOS: 7494
SOFTWARE: FastSeq for Windows, Version 4.0
SEQ ID NO 4971
LENGTH: 100
TYPE: PRT
ORGANISM: HUMAN
US-60-196-718-4971

Query Match 100.0%; Score 62; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
|||||
DB 45 TRSSGSIASNYQ 57

RESULT 8
PCT-US00-26524B-6775
Sequence 6775, Application PC/TUS0026524B
GENERAL INFORMATION:
APPLICANT: Birse et. al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PM005PCT
CURRENT APPLICATION NUMBER: PCT/US00/26524B
CURRENT FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6775
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (149)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-26524B-6775

Query Match 100.0%; Score 62; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 0.0028;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYQ 13
|||||
DB 44 TRSSGSIASNYQ 56

RESULT 9
PCT-US01-19110-1234
Sequence 1234, Application PC/TUS0119110
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523PCT
CURRENT APPLICATION NUMBER: PCT/US01/19110
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1234
LENGTH: 258
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-19110-1234

Query Match 90.3%; Score 56; DB 1; Length 258;
Best Local Similarity 92.3%; Pred. No. 0.066;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYQ 13
|||||
DB 170 TRSSGSIASNYQ 182

RESULT 10
US-09-880-748-1234
Sequence 1234, Application US/09880748
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1234
LENGTH: 258
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-1234

Query Match 90.3%; Score 56; DB 22; Length 258;

Best Local Similarity 92.3%; Pred. No. 0.066;
Matches 12: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVO 13
Db 170 TRSSGSIASNYVO 182

RESULT 11

US-09-043-522A-11
; Sequence 11, Application US/09043522A
; GENERAL INFORMATION:
; APPLICANT: Intracel Corporation
; APPLICANT: Pilkington, Glenn
; APPLICANT: Gilmour, Page
; TITLE OF INVENTION: Neutralizing Monoclonal Antibodies to Respiratory Syncytial Virus
; FILE REFERENCE: 58138-083
; CURRENT APPLICATION NUMBER: US/09/043.522A
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: region
; LOCATION: (1)..(13)
US-09-043-522A-11

Query Match 88.7%; Score 55; DB 14; Length 13;
Best Local Similarity 84.6%; Pred. No. 0.0032;
Matches 11: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVO 13
Db 1 TRAGSGSIASNYVO 13

RESULT 12

US-09-043-530-11
; Sequence 11, Application US/09043530
; GENERAL INFORMATION:
; APPLICANT: Intracel Corporation
; APPLICANT: Pilkington, Glenn
; APPLICANT: Gilmour, Page
; APPLICANT: Chanock, Robert
; APPLICANT: Crowe, James
; APPLICANT: Murphy, Brian
; TITLE OF INVENTION: Neutralizing Monoclonal Antibodies to Respiratory Syncytial Virus
; FILE REFERENCE: 58138-084
; CURRENT APPLICATION NUMBER: US/09/043.530
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: region
; LOCATION: (1)..(13)
US-09-043-530-11

Query Match 88.7%; Score 55; DB 14; Length 13;
Best Local Similarity 84.6%; Pred. No. 0.0032;
Matches 11: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVO 13
Db 1 TRAGSGSIASNYVO 13

RESULT 13
US-60-163-062-790
; Sequence 790, Application US/60163062
; GENERAL INFORMATION:
; APPLICANT: Bonazzoli, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CLO00134
; CURRENT APPLICATION NUMBER: US/60/163.062
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1302
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 790
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Human
US-60-163-062-790

Query Match 88.7%; Score 55; DB 24; Length 79;
Best Local Similarity 92.3%; Pred. No. 0.026;
Matches 12: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVO 13
Db 9 TGSSGSIASNYVO 21

RESULT 14

US-60-163-062-948
; Sequence 948, Application US/60163062
; GENERAL INFORMATION:
; APPLICANT: Bonazzoli, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CLO00134
; CURRENT APPLICATION NUMBER: US/60/163.062
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1302
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 948
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Human
US-60-163-062-948

Query Match 88.7%; Score 55; DB 24; Length 79;
Best Local Similarity 92.3%; Pred. No. 0.026;
Matches 12: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYVO 13
Db 9 TGSSGSIASNYVO 21

RESULT 15

US-60-163-123-1181
; Sequence 1181, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: Bonazzoli, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND USES
; FILE REFERENCE: CLO00137
; CURRENT APPLICATION NUMBER: US/60/163.123
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1986
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 1181
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Human
US-60-163-123-1181

```

Query Match	88.7%	Score 55:	DB 24:	Length 94:
Best Local Similarity	92.3%	Pred. NO. 0.032:		
Matches 12:	Conservative	0:	Mismatches 1:	Indels 0:
Gaps 0:				
Qy 1	TRSSGSINSNYQ	13		
Db 25	TGSSGSINSNYQ	37		

Search completed: November 19, 2001, 08:23:16
Job time: 729 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:31 ; Search time 11.44 Seconds
(without alignments)
6.053 Million cell updates/sec

Title: US-09-610-118-64

Perfect score: 62

Sequence: 1 TRSSGSIASNYVO 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 25229 seqs, 5326477 residues

Total number of hits satisfying chosen parameters: 25229

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	88.7	13	5	US-09-972-656-2
2	55	88.7	216	5	US-09-972-656-108
3	54	87.1	13	5	US-09-972-656-3
4	54	87.1	216	5	US-09-972-656-90
5	50	80.6	13	5	US-09-972-656-1
6	50	80.6	103	5	US-09-972-656-130
7	50	80.6	217	5	US-09-972-656-88
8	47	75.8	13	5	US-09-972-656-9
9	47	75.8	218	5	US-09-972-656-102
10	35	56.5	174	5	US-09-975-901-18
11	35	56.5	181	5	US-09-975-901-5
12	35	56.5	386	5	US-09-975-901-4
13	35	56.5	488	6	US-60-324-050-2
14	34	54.8	127	5	US-09-976-594-742
15	34	54.8	738	5	US-09-897-516-7782
16	33	53.2	257	5	US-09-897-516-7427
17	33	53.2	368	5	US-09-897-516-5907
18	33	53.2	381	5	US-09-815-242-5153
19	33	53.2	449	5	US-09-897-516-8080
20	32	51.6	173	5	US-09-787-126-10
21	32	51.6	173	5	US-09-787-126-12
22	32	51.6	173	5	US-09-787-126-18
23	32	51.6	182	5	US-09-787-126-16
24	32	51.6	187	5	US-09-787-126-8
25	32	51.6	188	5	US-09-787-126-14
26	32	51.6	316	5	US-09-787-126-4
27	32	51.6	316	5	US-09-787-126-6

28	32	51.6	316	5	US-09-957-944-8	Sequence 8, Appli
29	32	51.6	316	6	US-60-329-393-19	Sequence 19, Appli
30	32	51.6	316	6	US-60-329-360-6	Sequence 6, Appli
31	31	50.0	312	5	US-09-886-055-187	Sequence 187, App
32	31	50.0	317	5	US-09-787-126-2	Sequence 2, Appli
33	31	50.0	317	5	US-09-957-944-6	Sequence 6, Appli
34	31	50.0	570	5	US-09-877-633-13	Sequence 13, Appli
35	31	50.0	570	5	US-09-877-633-14	Sequence 14, Appli
36	31	50.0	652	5	US-09-815-242-13317	Sequence 1317, A
37	31	50.0	652	5	US-09-815-242-13673	Sequence 13673, A
38	31	50.0	3739	5	US-09-657-440-2	Sequence 2, Appli
39	30	48.4	772	5	US-09-815-242-13220	Sequence 13220, A
40	30	48.4	188	5	US-09-800-198-95	Sequence 95, Appli
41	30	48.4	272	5	US-09-815-242-10833	Sequence 10833, A
42	30	48.4	283	5	US-09-815-242-11268	Sequence 11268, A
43	30	48.4	300	5	US-09-800-198-25	Sequence 25, Appli
44	30	48.4	347	5	US-09-897-516-7301	Sequence 7301, Ap
45	30	48.4	349	5	US-09-897-516-6779	Sequence 6779, Ap

ALIGNMENTS

RESULT 1
US-09-972-656-2
; Sequence 2, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972.656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-972-656-2

Query Match 88.7%; Score 55; DB 5; Length 13;
Best Local Similarity 92.3%; Pred. No. 6.4e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OR 1 TRSSGSIASNYVO 13
DB 1 TRSSGSIASNYVO 13

RESULT 2
US-09-972-656-108
; Sequence 108, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972.656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-972-656-108

Query Match 88.7%; Score 55; DB 5; Length 216;

Best Local Similarity 92.3%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYVO 13
| | | | | | | | | | | | | | | | | |
Db 23 TRSSGSIASNYVO 35

RESULT 3
US-09-972-656-3
; Sequence 3, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-3

Query Match 87.1%; Score 54; DB 5; Length 13;
Best Local Similarity 92.3%; Pred. No. 9.8e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVO 13
| | | | | | | | | | | | | | | | | |
Db 1 TRSSGSIASNYVO 13

RESULT 4
US-09-972-656-90
; Sequence 90, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-90

Query Match 87.1%; Score 54; DB 5; Length 216;
Best Local Similarity 92.3%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVO 13
| | | | | | | | | | | | | | | | | |
Db 23 TRSSGSIASNYVO 35

RESULT 5
US-09-972-656-1
; Sequence 1, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma

; TITLE OF INVENTION: Neutralizing Activity
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-1

Query Match 80.6%; Score 50; DB 5; Length 13;
Best Local Similarity 84.6%; Pred. No. 0.00052;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYVO 13
| | | | | | | | | | | | | | | | | |
Db 1 TRSSGSIASNYVO 13

RESULT 6
US-09-972-656-130
; Sequence 130, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE: Misc.
; NAME/KEY: (7)..(7)
; LOCATION: (7)..(7)
; OTHER INFORMATION: Unidentified
; NAME/KEY: Misc.
; LOCATION: (33)..(33)
; OTHER INFORMATION: Unidentified
; NAME/KEY: Misc.
; LOCATION: (34)..(34)
; OTHER INFORMATION: Unidentified
; NAME/KEY: Misc.
; LOCATION: (35)..(35)
; OTHER INFORMATION: Unidentified
; NAME/KEY: Misc.
; LOCATION: (36)..(36)
; OTHER INFORMATION: Unidentified
US-09-972-656-130

Query Match 80.6%; Score 50; DB 5; Length 103;
Best Local Similarity 76.5%; Pred. No. 0.0045;
Matches 13; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 TRSSGSIASNYVO 13
| | | | | | | | | | | | | | | | | |
Db 24 TRSSGSIASNYVO 40

RESULT 7
US-09-972-656-88
; Sequence 88, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei

```
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE OF INVENTION: Neutralizing Activity
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 88
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-88
```

```
Query Match      80.6%; Score 50; DB 5; Length 217;
Best Local Similarity 84.6%; Pred. No. 0.0099;
Matches 11: Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 TRSSGSIASNTVQ 13
    1 |||||1111111
Db 23 TGSSGSIASNTVQ 35
```

```
RESULT 8
US-09-972-656-9
; Sequence 9, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE OF INVENTION: Neutralizing Activity
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-9
```

```
Query Match      75.8%; Score 47; DB 5; Length 13;
Best Local Similarity 83.3%; Pred. No. 0.0018;
Matches 10: Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 TRSSGSIASNTV 12
    1 |||||1111111
Db 1 TGSSGSIASNTV 12
```

```
RESULT 9
US-09-972-656-102
; Sequence 102, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE OF INVENTION: Neutralizing Activity
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 102
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-102
```

Query Match 75.8%; Score 47; DB 5; Length 218;

```
Best Local Similarity 83.3%; Pred. No. 0.035;
Matches 10: Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 TRSSGSIASNTV 12
    1 |||||1111111
Db 23 TGSSGSIASNTV 34
```

```
RESULT 10
US-09-975-901-18
; Sequence 18, Application US/09975901
; GENERAL INFORMATION:
; APPLICANT: Davis, George E.
; APPLICANT: Bell, Scott E.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING CMG PROTEINS, CMG PROTEINS, AN
; FILE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: 12740,0245.NPUS00
; CURRENT APPLICATION NUMBER: US/09/975,901
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,772
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-975-901-18
```

```
Query Match      56.5%; Score 35; DB 5; Length 174;
Best Local Similarity 50.0%; Pred. No. 4.3;
Matches 5: Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 SCGISASNTVQ 13
    1 |||||1111111
Db 19 SGSVANNWIE 28
```

```
RESULT 11
US-09-975-901-5
; Sequence 5, Application US/09975901
; GENERAL INFORMATION:
; APPLICANT: Davis, George E.
; APPLICANT: Bell, Scott E.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING CMG PROTEINS, CMG PROTEINS, AN
; FILE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: 12740,0245.NPUS00
; CURRENT APPLICATION NUMBER: US/09/975,901
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,772
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-975-901-5
```

```
Query Match      56.5%; Score 35; DB 5; Length 181;
Best Local Similarity 50.0%; Pred. No. 4.5;
Matches 5: Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 SGISASNTVQ 13
    1 |||||1111111
Db 19 SGSVANNWIE 28
```

RESULT 12
US-09-975-901-4
; Sequence 4, Application US/09975901

```
; GENERAL INFORMATION:
; APPLICANT: Davis, George E.
; APPLICANT: Bell, Scott E.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING CMG PROTEINS, CMG PROTEINS, AND M
; TITLE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: 12740.0245.NPUS00
; CURRENT APPLICATION NUMBER: US/09/975,901
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,772
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-975-901-4
```

```
Query Match          56.5%; Score 35; DB 5; Length 386;
Best Local Similarity 50.0%; Pred. No. 9.8;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 SCSIASNYQ 13
| | | | | | | | | |
Db 52 SGSVANNWIE 61
```

```
RESULT 13
US-60-324-050-2
; Sequence 2, Application US/603224050
; GENERAL INFORMATION:
; APPLICANT: Gene Logic, Inc.
; APPLICANT: Nation, Michelle
; APPLICANT: Diggans, James C.
; APPLICANT: Lu, Sun
; APPLICANT: Porter, Mark
; TITLE OF INVENTION: Genes Associated with Cervical and Uterine Cancer
; FILE REFERENCE: 44921-5093-PR
; CURRENT APPLICATION NUMBER: US/60/324,050
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-324-050-2
```

```
Query Match          56.5%; Score 35; DB 6; Length 488;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 SCSIASNYQ 13
| | | | | | | | | |
Db 52 SGSVANNWIE 61
```

```
RESULT 14
US-09-976-594-742
; Sequence 742, Application US/09976594
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
```

```
; SEQ ID NO 742
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 2847188C01
US-09-976-594-742
```

```
Query Match          54.8%; Score 34; DB 5; Length 127;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 TRSSGSIASNYV 12
| | | | | | | | | |
Db 98 SRSQGNVKSNEV 109
```

```
RESULT 15
US-09-897-516-7782
; Sequence 7782, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huebner, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 7782
; LENGTH: 758
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-7782
```

```
Query Match          54.8%; Score 34; DB 5; Length 758;
Best Local Similarity 61.5%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 TRSSGSIASNYV 13
| | | | | | | | | |
Db 268 TYSSINIASNYQ 280
```

Search completed: November 19, 2001, 08:23:32
Job time: 744 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:13:25 ; Search time 43.5 Seconds
(without alignments)
6.725 Million cell updates/sec

Title: US-09-610-118-64
Perfect score: 62
Sequence: 1 TRSSGSIASNYVO 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCrus.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfilltest.pep:*

Result No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	62.9	13	1	US-07-988-925-4
2	39	62.9	13	2	US-08-362-780-4
3	39	62.9	110	1	US-07-988-925-16
4	39	62.9	110	2	US-08-362-780-16
5	39	62.9	110	2	US-08-362-780-26
6	38	61.3	505	5	PCT-US95-05008-17
7	35	56.5	132	2	US-08-649-991-91
8	35	56.5	132	2	US-08-649-991-92
9	35	56.5	512	2	US-08-649-991-16
10	34	54.8	29	4	US-09-070-504-1
11	34	54.8	30	4	US-09-070-504-2
12	34	54.8	37	1	US-08-233-389C-9
13	34	54.8	37	2	US-08-801-863-9
14	34	54.8	37	2	US-08-486-596A-9
15	34	54.8	37	2	US-09-004-713-9
16	34	54.8	37	2	US-08-619-841-2
17	34	54.8	37	2	US-08-446-929A-2
18	34	54.8	37	4	US-09-070-504-3
19	34	54.8	37	4	US-09-070-504-8
20	34	54.8	37	4	US-09-070-504-9
21	34	54.8	37	6	US-07-776-272-20
22	34	54.8	38	1	US-08-211-202-111
23	34	54.8	98	1	US-08-665-202-38
24	34	54.8	109	1	US-08-276-852-147
25	34	54.8	109	1	US-08-899-575-147
26	34	54.8	109	1	US-08-899-575-147
27	34	54.8	109	1	US-08-899-575-147

28	34	54.8	109	4	US-09-240-274-59	Sequence 59, Appl
29	34	54.8	109	4	US-09-025-769B-32	Sequence 32, Appl
30	34	54.8	109	4	US-09-025-769B-51	Sequence 51, Appl
31	34	54.8	109	5	PCT-US95-08743-147	Sequence 60, Appl
32	34	54.8	110	4	US-09-240-274-60	Sequence 21, Appl
33	34	54.8	111	4	US-08-983-607-21	Sequence 23, Appl
34	34	54.8	111	4	US-08-983-607-23	Sequence 113, App
35	34	54.8	113	1	US-08-211-202-113	Sequence 10, Appl
36	34	54.8	113	1	US-08-211-202-114	Sequence 11, Appl
37	33	53.2	37	4	US-09-070-504-10	Sequence 15, Appl
38	33	53.2	37	4	US-09-070-504-11	Sequence 23, Appl
39	33	53.2	142	2	US-08-480-774A-2	Sequence 5, Appl
40	33	53.2	187	1	US-08-403-378B-15	Sequence 2, Appl
41	33	53.2	332	4	US-09-331-581-23	Sequence 5, Appl
42	43	53.2	431	2	US-08-665-647-5	Sequence 2, Appl
43	33	53.2	588	1	US-08-391-615-5	Sequence 2, Appl
44	33	53.2	691	1	US-08-178-477B-2	Sequence 2, Appl
45	33	53.2	1196	4	US-08-881-706-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-07-988-925-4
Sequence 4, Application US/07988925
Patent No. 5585097
GENERAL INFORMATION:
APPLICANT: Bolt, Sarah L
APPLICANT: Clark, Michael R
APPLICANT: Gorman, Scott D
APPLICANT: Roulledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: antibody preparation
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Nixon and Vanderhye pc
STREET: 11th Floor, 1100 No. 5585097th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988, 925
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206422.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CB92/01933
FILING DATE: 21-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Michard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164100
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-988-925-4

Query Match 62.9%; Score 39; DB 1; Length 13;
Best Local Similarity 66.7%; Pred. No. 0.26;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYV 12
| | | | | : | | | |
Db 1 TLSSGNIENNYV 12

RESULT 2

US-08-362-780-4
Sequence 4, Application US/08362780
Patent No. 5968509
GENERAL INFORMATION:
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: Antibody Preparation
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye PC
STREET: 8th Floor, 1100 No. 5968509th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,780
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/862,543
FILING DATE: 23-JUNE-1992
APPLICATION NUMBER: GB 9021679.7
FILING DATE: 05-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB91/01726
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mitchell, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-780-4

Query Match 62.9%; Score 39; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 0.26;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYV 12
| | | | | : | | | |
Db 1 TLSSGNIENNYV 12

RESULT 3

US-07-988-925-16
Sequence 16, Application US/07988925
Patent No. 5585097
GENERAL INFORMATION:

APPLICANT: Bolt, Sarah L
APPLICANT: Clark, Michael R
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: antibody preparation
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye PC
STREET: 11th Floor, 1100 No. 5585097th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988,925
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206422.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB92/01933
FILING DATE: 21-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mitchell, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-988-925-16

Query Match 62.9%; Score 39; DB 1; Length 110;
Best Local Similarity 66.7%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TRSSGSIASNYV 12
| | | | | : | | | |
Db 23 TLSSGNIENNYV 34

RESULT 4

US-08-362-780-16
Sequence 16, Application US/08362780
Patent No. 5968509
GENERAL INFORMATION:
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: Antibody Preparation
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye PC
STREET: 8th Floor, 1100 No. 5968509th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,780
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/862,543
FILING DATE: 23-JUNE-1992
APPLICATION NUMBER: GB 9021679.7
FILING DATE: 05-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB91/01726
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-780-16

Query Match 62.9%; Score 39; DB 2; Length 110;
Best Local Similarity 66.7%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
| | | | | : | | | |
Db 23 TLSSGNIENNYV 34

RESULT 5
US-08-362-780-26
Sequence 26, Application US/08362780
Patent No. 5968509
GENERAL INFORMATION:
APPLICANT: Gorman, Scott D
APPLICANT: Routledge, Edward G
APPLICANT: Waldmann, Herman
TITLE OF INVENTION: Antibody Preparation
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye PC
STREET: 8th floor, 1100 No. 5968509th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,780
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/862,543
FILING DATE: 23-JUNE-1992
APPLICATION NUMBER: GB 9021679.7
FILING DATE: 05-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB91/01726

FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-780-26

Query Match 62.9%; Score 39; DB 2; Length 110;
Best Local Similarity 66.7%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TRSSGSIASNYV 12
| | | | | : | | | |
Db 23 TLSSGNIENNYV 34

RESULT 6
PCT-US95-05008-17
Sequence 17, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid

STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-05008-17

Query Match 61.3%; Score 38; DB 5; Length 505;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TRSSGSIASNY 12
DB 101 TRREGYIPSNY 112

RESULT 7

US-08-649-991-91
Sequence 91, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7176
TELEFAX: 202-467-7000
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-649-991-91

Query Match 56.5%; Score 35; DB 2; Length 132;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRSSGSIASNY 11
DB 122 TRSSQVSQSNY 132

RESULT 8

US-08-649-991-92
Sequence 92, Application US/08649991
Patent No. 5919462
GENERAL INFORMATION:
APPLICANT: Narwa, Remy
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS DERIVED FROM THE
TITLE OF INVENTION: HIV-1 VIRUS GENOME, CORRESPONDING PEPTIDES AND THEIR
TITLE OF INVENTION: APPLICATIONS AS REAGENTS FOR EVALUATION OF THE RISK OF
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,991
FILING DATE: 17-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9505914
FILING DATE: 18-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7176
TELEFAX: 202-467-7000
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-649-991-92

Query Match 56.5%; Score 35; DB 2; Length 132;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRSSGSIASNY 11
DB 122 TRSSQVSQSNY 132

RESULT 9
PCT-US95-05008-16
Sequence 16, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds

```

1 NAME: McCormack, Myra H
2 REGISTRATION NUMBER: 36,602
3 REFERENCE/DOCKET NUMBER: 180.00020101
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: 612/305-1220
6 TELEFAX: 612/305-1228
7 INFORMATION FOR SEQ ID NO: 1:
8 SEQUENCE CHARACTERISTICS:
9 LENGTH: 29 amino acids
10 TYPE: amino acid
11 STRANDEDNESS: single
12 TOPOLOGY: linear
13 MOLECULE TYPE: peptide
14 US-09-070-504-1
15
16 Query Match 54.8%; Score 34; DB 4; Length 29;
17 Best Local Similarity 50.0%; Pred. No. 5.6;
18 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0
19
20 Oy 1 TRSSGSISANY 12
21 :111:11:1
22 Db 9 SRSGGVKSNFY 20
23
24 RESULT 11
25 US-09-070-504-2
26 Sequence 2, Application US/09070504
27 Patent No. 6268474
28
29 GENERAL INFORMATION:
30 APPLICANT: Smith, Derek D.
31 APPLICANT: Saha, Shankar
32 APPLICANT: Abel, Peter W.
33 TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF CGRP-RECEPTOR
34 NUMBER OF SEQUENCES: 23
35 CORRESPONDENCE ADDRESS:
36 ADDRESSEE: Metting, Raasch & Gebhardt, P.A.
37 STREET: 119 No. 6268474th Fourth Street
38 CITY: Minneapolis
39 STATE: MN
40 COUNTRY: USA
41 ZIP: 55401
42
43 COMPUTER READABLE FORM:
44 MEDIUM TYPE: Floppy disk
45 COMPUTER: IBM PC compatible
46 OPERATING SYSTEM: PC-DOS/MS-DOS
47 SOFTWARE: Patentin Release #1.0, Version #1.30
48 CURRENT APPLICATION DATA:
49 APPLICATION NUMBER: US/09/070,504
50 FILING DATE: 30-APR-1998
51 CLASSIFICATION:
52 ATTORNEY/AGENT INFORMATION:
53 NAME: McCormack, Myra H
54 REGISTRATION NUMBER: 36,602
55 REFERENCE/DOCKET NUMBER: 180.00020101
56 TELECOMMUNICATION INFORMATION:
57 TELEPHONE: 612/305-1220
58 TELEFAX: 612/305-1228
59 INFORMATION FOR SEQ ID NO: 2:
60 SEQUENCE CHARACTERISTICS:
61 LENGTH: 30 amino acids
62 TYPE: amino acid
63 STRANDEDNESS: single
64 TOPOLOGY: linear
65 MOLECULE TYPE: peptide
66 US-09-070-504-2
67
68 Query Match 54.8%; Score 34; DB 4; Length 30;
69 Best Local Similarity 50.0%; Pred. No. 5.8;
70 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 TRSSGSIASNYV 12
:111:11:1
DB 9 SRSGMWKSNFV 20

RESULT 12

US-08-233-389C-9
; Sequence 9, Application US/08233389C
; Patent No. 5639855
; GENERAL INFORMATION:
; APPLICANT: KITAMURA, Kazuo
; APPLICANT: KANGAWA, Kenji
; APPLICANT: MATSUO, Hisayuki
; APPLICANT: ETO, Tanenao
; TITLE OF INVENTION: ADRENOMEDULLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10020

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/233.389C
; FILING DATE: 26-APR-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY JR., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-233-389C-9

Query Match 54.8%; Score 34; DB 1; Length 37;
Best Local Similarity 50.0%; Pred. No. 7.4;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYV 12
:111:11:1
DB 17 SRSGMWKSNFV 28

RESULT 13

US-08-801-863-9
; Sequence 9, Application US/08801863
; Patent No. 5830703
; GENERAL INFORMATION:
; APPLICANT: KITAMURA, Kazuo
; APPLICANT: KANGAWA, Kenji
; APPLICANT: MATSUO, Hisayuki
; APPLICANT: ETO, Tanenao
; TITLE OF INVENTION: ADRENOMEDULLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY

COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.863
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY JR., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-5 DIV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-801-863-9

Query Match 54.8%; Score 34; DB 2; Length 37;
Best Local Similarity 50.0%; Pred. No. 7.4;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 TRSSGSIASNYV 12
:111:11:1
DB 17 SRSGMWKSNFV 28

RESULT 14

US-08-486-596A-9
; Sequence 9, Application US/08486596A
; Patent No. 5837823
; GENERAL INFORMATION:
; APPLICANT: KITAMURA, Kazuo
; APPLICANT: KANGAWA, Kenji
; APPLICANT: MATSUO, Hisayuki
; APPLICANT: ETO, Tanenao
; TITLE OF INVENTION: ADRENOMEDULLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10020
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486.596A
; FILING DATE: JUNE 7, 1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY JR., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-5 DIV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-486-596A-9

Query Match 54.8%; Score 34; DB 2; Length 37;
Best Local Similarity 50.0%; Pred. No. 7.4;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYV 12
: | | : | | :
Db 17 SRSGGMVKSNEV 28

RESULT 15
US-09-004-713-9
; Sequence 9, Application US/09004713
; Patent No. 5910416
; GENERAL INFORMATION:
; APPLICANT: KITAMURA, Kazuo
; APPLICANT: KANGAWA, Kenji
; APPLICANT: MATSUO, Hisayuki
; APPLICANT: ETO, Tanenao
; TITLE OF INVENTION: ADRENOMEDULLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C/O FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,713
; FILING DATE: JANUARY 7, 1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY JR., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHCN-5 DIV2 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-004-713-9

Query Match 54.8%; Score 34; DB 2; Length 37;
Best Local Similarity 50.0%; Pred. No. 7.4;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 TRSSGSIASNYV 12
: | | : | | :
Db 17 SRSGGMVKSNEV 28

Search completed: November 19, 2001, 08:13:25
Job time: 138 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:12:35 ; Search time 77.92 Seconds
(without alignments)
5.446 Million cell updates/sec

Title: US-09-610-118-65
Perfect score: 38
Sequence: 1 EDNRPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	22	AAB61297
2	34	89.5	7	18	AAW16592
3	34	89.5	106	15	AAW54054
4	34	89.5	106	17	AAW01529
5	34	89.5	106	18	AAW24991
6	34	89.5	109	18	AAW16588
7	34	89.5	109	18	AAW14785
8	34	89.5	114	21	AAW53654
9	33	86.8	112	18	AAW08488
10	32	84.2	108	15	AAW54051
11	32	84.2	108	17	AAW01523

12	32	84.2	108	18	AAW24988	Monoclonal antibody
13	31	81.6	108	17	AAW88719	Human antibody 1am
14	31	81.6	108	21	AAW68898	The variable region of
15	31	81.6	108	21	AAW68900	Variable region of
16	30	78.9	109	18	AAW27547	Human Ab light chain
17	30	78.9	151	21	AAW42871	Arbidopsis thalia
18	30	78.9	155	15	AAW45451	Adenovirus E1A-F P
19	30	78.9	158	19	AAW85874	S. pneumoniae deri
20	30	78.9	161	21	AAW42870	Arbidopsis thalia
21	30	78.9	232	21	AAW42869	Arbidopsis thalia
22	30	78.9	462	17	AAW00167	E1AF matrix metallo
23	30	78.9	462	17	AAW44725	Human polyomavirus
24	30	78.9	555	19	AAW49010	Polyomavirus PEA3
25	30	78.9	555	19	AAW44724	Murine polyomaviru
26	30	78.9	555	22	AAW60290	Mouse PEA-3 protei
27	30	78.9	620	17	AAW94535	ITK tyrosine kinas
28	30	78.9	858	22	AAW71986	Rat autotaxin prot
29	30	78.9	858	22	AAW71995	Rat autotaxin vari
30	30	78.9	858	22	AAW71996	Rat autotaxin vari
31	30	78.9	858	22	AAW71997	Rat autotaxin vari
32	30	78.9	858	22	AAW71998	Rat autotaxin vari
33	30	78.9	858	22	AAW71999	Rat autotaxin vari
34	30	78.9	885	22	AAW71989	Rat brain autotaxi
35	29	76.3	98	21	AAW40143	Anti-hu112 antibody
36	29	76.3	109	19	AAW54018	Human Anti-CD4 ant
37	29	76.3	111	18	AAW13528	Anti-melanoma ligh
38	29	76.3	112	18	AAW13523	Anti-melanoma ligh
39	29	76.3	129	18	AAW34081	Human monoclonal a
40	29	76.3	132	18	AAW16323	Human ARF-p19, a n
41	29	76.3	132	21	AAW79145	Human ARF-p19 cell
42	29	76.3	202	8	AAW70319	Protein M2 encoded
43	29	76.3	203	21	AAW54264	Human pancreatic c
44	29	76.3	258	21	AAW47237	Arbidopsis thalia
45	29	76.3	260	21	AAW47236	Arbidopsis thalia

ALIGNMENTS

RESULT 1	
AAW61297	standard; Peptide: 7 AA.
XX	
AC	AAW61297;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Anti-TANGO 268 scfv CDR, SEQ ID NO: 65.
XX	
KW	Human; antibody; scfv; CDR; complementarity determining region;
KW	TANGO 268; cardiant; cerebroprotective; cytosolic; anticoagulant;
KW	thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;
KW	platelet membrane glycoprotein receptor; bleeding disorder;
KW	blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;
KW	ischaemia; cardiovascular disease; immunological disease; liver disorder;
KW	cancer.
XX	
OS	Homo sapiens.
XX	
PN	WO200100810-A1.
XX	
PD	04-JAN-2001.
XX	
PF	30-JUN-2000; 2000WO-US18152.
XX	
PR	30-JUN-1999; 99US-0345468.
XX	
PR	06-DEC-1999; 99US-0454824.
XX	
PR	14-FEB-2000; 2000US-0503387.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Busfield SJ, Villette J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI	Qian MD, Kingsbury G;

SQ Sequence 7 AA;

QY	1	EDNQrps	7
nb	1	ednqrps	7

RESULT	2
AAW16592	
ID	AAW16592 standard; Peptide; 7 AA

AC	AAW16592;
XX	
DT	30-NOV-1997 (first entry)
XX	

Anti-RSV F glycoprotein antibody RSVF2-5 light chain VL6 CDR2.

KW Respiratory syncytial virus; RSV; monoclonal antibody; CDR;
KW complementarity determining region; pneumonia; bronchiolitis;
KW diagnosis; therapy; vaccine: RSVF2-5.

Homo sapiens.

PN W09710846-A1.

PD 27-MAR-1997.

PF 18-SEP-1996; 96WO-US14937.

PR 18-SEP-1995; 95US-0003931.

PA (INTR-) INTRACEL CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES
XX

PI Chanock RM, Crowe JE, Gilmour PS, Murphy BR, Pilkington GR,
XX
DR WPI, 1997-202621/18.

PT Composition comprising respiratory syncytial virus antibody - useful
for treatment or prophylaxis of active disease or infection
XX
PS Claim 8; Page 56; 71pp: English

SQ Sequence 7 AA;

```
Query Match      89.5%; Score 34; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	1	EDNQRP	6
Db	1	ednqrp	6

```

RESULT      3
AAR54054
ID  AAR54054 standard; Protein; 106 AA

```

AC	AAR54054;	
XX		
DT	08-NOV-1994	(first entry)

DE Sequence of the VL region of monoclonal antibody MD3-4 against
DE hepatitis B virus surface antigen.

KW Hepatitis B virus; surface antigen; monoclonal antibody; therapy;
 HBsAg; diagnosis; HBV.

OS Synthetic.

Key	Location/Qualifiers
EH	1..95
EH	/label= V III
ET	24..33
ET	/label= CDR 1
ET	50..56
ET	/label= CDR 2
ET	88..96
ET	/label= CDR 3
ET	95..106
ET	/label= j 3

PN W09411495-A

PD 26-MAY-1994

PF 06-NOV-1992; 92WO-US09749.
YY

PR 06-NOV-1992; 92W0-US09749.
YY

PA (SANO) SANDOZ LTD
VY

PI Ostberg LG;

WPI; 1994-183497/22

```

DR N-PSDB; AAO64057.
XX
XX Monoclonal antibodies active against Hepatitis B surface antigen
PT - for diagnosis and treatment of Hepatitis B virus
XX
XX Example; Page 43; 53pp; English.
XX
XX Human volunteers were immunised with hepatitis B vaccine. MD3-4, ZM1-
CC 2, ZM1-1 and PE1-1 hybridoma cell lines were derived from
CC lymphocytes of individuals immunised with Hepavax (Merck & Co).
CC Antibodies PE1-1, ZM1-1, ZM1-2 and MD3-4 belong to the IgG1 clas.
CC The cell lines producing PE1-1, ZM1-1 and ZM1-2 were deposited as
CC ATCC HB9234, 9191 and 9192 respectively. The cell lines all behave
CC as typical (mouse x human) x human hybridomas and produce their
CC respective Abs in concs. ranging up to 25 mg/l in standard
CC suspension culture. The heavy variable (VH) and light variable (LH)
CC chains of Abs PE1-1, ZM1-1, ZM1-2 and MD3-4 were isolated and
CC sequenced. Total RNA was extracted from 10(7) hybridoma cells
CC of each cell line. ss DNA was synthesised using AMV-reverse
CC transcriptase and oligo-dt as primer. PCRs were performed and
CC amplified DNA was size selected. ss DNA for sequencing was isolated
CC from each positive clone after superinfection with M13K07.
CC Sequencing was by the dideoxy chain termination method (Sanger
CC et al.).
XX
XX Sequence 106 AA:
SQ
Query Match 89.5%; Score 34; DB 15; length 106;
Best Local Similarity 85.7%; Pred. No. 8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORPS 7
Db 49 ednkrrps 55
|||:|
RESULT 4
AAW01529
ID AAW01529 standard; Protein; 106 AA.
XX
AC AAW01529;
XX
DT 04-MAR-1997 (first entry)
XX
DE Monoclonal antibody MD3-4 light chain variable region.
XX
XX Monoclonal antibody; diagnosis; treatment; infection; hepatitis B;
KW xenogeneic hybridoma; SPAZ 4; PE1-1; ZM1-2; MD3-4; L03-3;
KW IgG1 class; heavy chain; light chain; variable region.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 1..94
FT /label= V_kappa_III_region
FT Region 24..33
FT /label= CDR1
FT Region 34..49
FT /label= framework_region
FT Region 50..56
FT /label= CDR2
FT Region 57..87
FT /label= framework_region
FT Region 88..97
FT /label= CDR3
FT Region 95..106
FT /label= J_kappa_3_region
XX
XX USS565354-A.
XX
XX 15-OCT-1996.
XX

```

```

PF 05-SEP-1986; 86US-0904517.
XX
XX 21-APR-1992; 92US-0871426.
PR 05-SEP-1986; 86US-0904517.
PR 31-OCT-1986; 86US-0925196.
PR 11-MAY-1988; 88US-0192754.
PR 15-JUN-1990; 90US-0538796.
PR 27-MAR-1991; 91US-0676036.
PR 14-JUN-1994; 94US-0259372.
XX
XX (SANO ) SANDOZ LTD.
XX
XX Ostberg LG;
XX
XX WPI; 1996-476304/47.
DR N-PSDB; AAT46135.
XX
XX Human monoclonal antibodies specific for hepatitis B surface antigen
PT - are used to treat or prevent infection or in diagnostic assays
XX
XX Example 9; Column 43-46; 26pp; English.
XX
XX Monoclonal antibodies effective for the diagnosis and treatment of
CC diseases caused by infection with hepatitis B have been prepared from a
CC cell line obtained by fusing a xenogeneic hybridoma designated SPAZ 4
CC with blood cells of a patient immunised with hepatitis B vaccine.
CC Specific antibodies are PE1-1, ZM1-1, ZM1-2, MD3-4 and L03-3, each of
CC these being of the IgG1 class. The present sequence is the light
CC variable chain of MD3-4.
XX
XX Sequence 106 AA:
SQ
Query Match 89.5%; Score 34; DB 17; length 106;
Best Local Similarity 85.7%; Pred. No. 8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORPS 7
Db 49 ednkrrps 55
|||:|
RESULT 5
AAW24991
ID AAW24991 standard; Protein; 106 AA.
XX
AC AAW24991;
XX
DT 08-DEC-1997 (first entry)
XX
DE Monoclonal antibody MD3-4 VL region.
XX
XX Heavy chain; light chain; variable region; human; monoclonal antibody;
KW immunisation; hepatitis B virus; HBV; vaccine; mouse; fusion; xenogeneic;
KW peripheral blood lymphocyte; surface antigen; cell culture; ion exchange;
KW chromatography; size separation; primer; PCR; polymerase chain reaction;
KW amplification; hybridoma; infection; immunosuppression; hepatitis;
KW liver transplant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 1..94
FT /note= "VIII region"
FT Domain 24..33
FT /note= "complementarity determining region 1"
FT Domain 50..56
FT /note= "complementarity determining region 2"
FT Domain 88..96
FT /note= "complementarity determining region 3"
FT Region 95..106
FT /note= "J13 region"
XX

```

PN US5648077-A.
 XX
 PD 15-JUL-1997.
 XX
 XX
 PE 05-SEP-1986; 86US-0904517.
 XX
 PR 21-APR-1992; 92US-0871426.
 PR 05-SEP-1986; 86US-0904517.
 PR 31-OCT-1986; 86US-0925196.
 PR 11-MAY-1988; 88US-0192754.
 PR 15-JUN-1990; 90US-0538796.
 PR 27-MAR-1991; 91US-0676036.
 PR 14-JUN-1994; 94US-0259372.
 PR 06-JUN-1995; 95US-0468671.
 XX
 PA (SANO) SANDOZ LTD.
 XX
 PI Ostberg LG;
 XX
 DR WPI; 1997-372021/34.
 DR N-PSDB; AAT85845.
 XX
 PT Treatment of hepatitis B - with human monoclonal antibody
 PT
 PS Example 8; Column 23-24; 25pp; English.
 XX
 CC This is the amino acid sequence of the light chain variable (VL) region
 CC from the human monoclonal antibody (Mab) MD3-4. The Mab was generated
 CC by immunising humans with a hepatitis B virus (HBV) vaccine, isolating
 CC peripheral blood lymphocytes (PBL) and fusing them with a mouse/human
 CC xenogeneic cell line SP2-4. 5 cell lines were isolated: PEI-1, ZM1-1,
 CC ZM1-2, MD3-4 and L03-3. The cell lines were then tested for production
 CC of an anti-hepatitis B virus surface antigen antibody by ELISA. The Mabs
 CC are then purified from large scale cell culture by protein A
 CC chromatography, size separation on Sephacryl S300 gel and ion exchange
 CC chromatography on Q-sepharose. The heavy and light chains of the Mabs
 CC were isolated and their amino acid sequences determined. Primers were
 CC generated and used to amplify cDNA synthesised from RNA purified from
 CC each hybridoma cell line. The sequences of the heavy and light chains
 CC (nucleic acid and amino acid) from Mabs PEI-1, ZM1-1, ZM1-2 and MD3-4
 CC are shown in AAT85838-45 and AAM24984-91. The Mabs can be used to treat
 CC HBV infections in immunosuppressed patients or patients with chronic
 CC active hepatitis, especially liver transplant patients.
 CC
 XX
 SQ Sequence 106 AA;
 Query Match 89.5%; Score 34; DB 18; Length 106;
 Best Local Similarity 85.7%; Pred. No. 8;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EDNORPS 7
 DB 49 ednkrps 55
 RESULT 6
 AAM16588
 ID AAM16588 standard; Protein: 109 AA.
 XX
 AC AAM16588;
 XX
 DT 30-NOV-1997 (first entry)
 XX
 DE Anti-RSV F glycoprotein antibody RSVF2-5 light chain VL6.
 XX
 KW Respiratory syncytial virus; RSV; monoclonal antibody; CDR;
 KW complementarity determining region; pneumonia; bronchiolitis;
 KW diagnosis; therapy; vaccine; RSVF2-5.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT Peptide 1..19
 FT /label= FR1
 FT /note= "framework region 1"
 FT Region 20..32
 FT /label= CDR1
 FT /note= "complementarity determining region 1
 FT (Claim 9)"
 FT Region 33..47
 FT /label= FR2
 FT /note= "framework region 2"
 FT Region 48..54
 FT /label= CDR2
 FT /note= "complementarity determining region 2
 FT (Claim 8)"
 FT Region 55..88
 FT /label= FR3
 FT /note= "framework region 3"
 FT Region 89..98
 FT /label= CDR3
 FT /note= "complementarity determining region 3
 FT (Claim 7)"
 FT Region 99..109
 FT /label= FR4
 FT /note= "framework region 4"
 XX
 PN WO9710846-A1.
 XX
 PD 27-MAR-1997.
 XX
 PF 18-SEP-1996; 96WO-US14937.
 XX
 PR 18-SEP-1995; 95US-0009331.
 XX
 PA (INTR-) INTRACEL CORP.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Chanock RM, Crowe JE, Gilmour PS, Murphy BR, Pilkington GR;
 DR WPI; 1997-202621/18.
 DR N-PSDB; AAT66557.
 PT Composition comprising respiratory syncytial virus antibody - useful
 PT for treatment or prophylaxis of active disease or infection
 PT
 PS Claim 10; Page 52-53; 71pp; English.
 XX
 CC This polypeptide sequence comprises the light chain VL6 region
 CC of a novel neutralising human monoclonal antibody to respiratory
 CC syncytial virus (RSV). The antibody, designated RSVF2-5 (ATCC
 CC 69909), selectively binds to an RSV F glycoprotein epitope. cDNA
 CC encoding the RSVF2-5 Fd VH3 and light chain VL6 (see AAT66556-57)
 CC was isolated from a phage library prepared from RNA extracted from
 CC the peripheral blood lymphocytes of an HIV-1 infected donor. A
 CC claimed pharmaceutical preparation comprises a carrier and an
 CC antibody that includes the RSVF2-5 VH3 CDR3 (AAM16584) and which may
 CC also include the VH3 CDR2 (AAM16584) and/or CDR1 or the entire Fd
 CC Fd region (AAM16580), or is an Fd fragment and further includes the
 CC RSVF2-5 VL6 CDR3 (AAM16594), CDR2 (AAM16592), CDR1 (AAM16590) or entire
 CC light chain (AAM16588). The preparation alternatively comprises a
 CC carrier and a vector that includes a nucleotide sequence encoding
 CC the antibody. The preparations can be used for the treatment or
 CC prophylaxis of active RSV disease or infection (claimed), and may
 CC also be used for RSV detection. The antibody binds and neutralises
 CC antigenic subgroups A and B of RSV with high efficiency.
 XX
 SQ Sequence 109 AA;
 Query Match 89.5%; Score 34; DB 18; Length 109;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EDNORP 6

Db 48 edngrp 53

```
RESULT 7
ID AAM14785 standard; Protein: 109 AA.
XX
XX AAM14785;
AC
XX
XX
DT 20-JUN-1997 (first entry)
DE Human monoclonal antibody RSVF2-5 light chain VL6.
XX
XX Respiratory syncytial virus; RSV; monoclonal antibody; Mab;
KM diagnosis; prophylaxis; immunotherapy; therapy; Fab; Fd.
XX
OS Homo sapiens.
FH
XX Key Location/Qualifiers
FT Region 1..19
FT Region /label= FRI
FT Region 20..32
FT Region /label= CDR1
FT Region /note= "Claim 9"
FT Region 33..47
FT Region /label= FR2
FT Region 48..54
FT Region /label= CDR2
FT Region /note= "Claim 8"
FT Region 55..88
FT Region /label= FR3
FT Region 89..98
FT Region /label= CDR3
FT Region /note= "Claim 7"
FT Region 99..109
FT Region /label= FR4
XX
XX W0971177-A1.
PN
XX
XX 27-MAR-1997.
PD
XX
XX 18-SEP-1996; 96WO-US14944.
PF
XX
XX 18-SEP-1995; 95US-0003931.
PR
XX
XX (INTR-) INTRACEL CORP.
PA
XX Gilmour PS, Pilkington GR;
PI
XX
XX WPI: 1997-202886/18.
DR N-PSDB; AAT63418.
DR
XX
XX Monoclonal antibody specific for respiratory syncytial virus - used
PT for diagnosis and immuno-prophylaxis or immuno-therapy of RSV
PS disease
XX
XX Claim 10; Page 49-50; 68pp; English.
XX
XX The light chain variable region (AAM14785) and heavy chain variable
CC region (AAM14784) are provided of a novel, fully human monoclonal
CC antibody (Mab), designated RSVF2-5, which selectively binds to
CC the F glycoprotein of respiratory syncytial virus (RSV) and which
CC neutralises RSV in vivo. To obtain RSVF2-5, a packaged phage
CC library was pred. from amplified heavy and light chain variable
CC region clones derived from the peripheral blood lymphocytes of an
CC HIV-1 infected donor. The phage library was panned with RSV
CC proteins and isolated clones were sequenced. Isolated nucleic
CC acids (AAT63417-18) can be used to produce novel polypeptides, esp.
CC Fd or Fab fragments useful in the diagnosis of RSV disease, in
CC methods for detecting the presence of RSV in a sample, and in the
XX immunoprophylaxis and immunotherapy of RSV disease.
```

SO Sequence 109 AA:

Query Match 89.5%; Score 34; DB 18; Length 109;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDNORP 6
Db 48 edngrp 53

```
RESULT 8
ID AAB53654 standard; Protein: 114 AA.
XX
XX AAB53654;
AC
XX
XX 09-MAR-2001 (first entry)
DT
XX
XX Human colon cancer antigen protein sequence SEQ ID NO:1194.
DE
XX
XX Human: colon cancer; colon cancer antigen; diagnosis; detection;
KM identification; cytosolic; cardioactive; neuroprotective; vulnereary;
KM immunomodulatory; muscular; gynaecological; gastrointestinal;
KM nephrotoxic; antiinfective; antibacterial; gene therapy; wound;
KM neural disorder; immune system disorder; muscular disorder;
KM reproductive disorder; gastrointestinal disorder; renal disorder;
KM infectious disease; cardiovascular disorder.
XX
XX Homo sapiens.
OS
XX
XX W0200055351-A1.
PN
XX
XX 21-SEP-2000.
PD
XX
XX 08-MAR-2000; 2000WO-US05883.
PF
XX
XX 12-MAR-1999; 99US-0124270.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Ruben SM;
PI
XX
XX WPI: 2000-587534/55.
DR
XX
XX N-PSDB: AAC98411.
DR
XX
XX Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PS disorders such as colon cancer -
XX
XX Claim 11; Page 1776; 2104pp; English.
XX
XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB55234 to AAB54006. The
CC human colon cancer antigens can have cytosolic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnereary, nephrotoxic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins
CC may also be used to prevent diseases such as neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, wounds, renal disorders, infectious
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC AAB54007 represent sequences used in the exemplification of the present
XX invention.
XX Sequence 114 AA;
```

Query Match 89.5%; Score 34; DB 21; Length 114;
 Best Local Similarity 85.7%; Pred. No. 8.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDNORPS 7
 111111
 Db 37 ednkrps 43

RESULT 9

AAW08488
 ID AAW08488 standard; Protein; 112 AA.

XX AC AAW08488;

XX DT 23-SEP-1997 (first entry)

XX DE C6 human sFv antibody C6.5 light chain shuffled mutant C6VLB.

XX KW Tumour; Immune response; cytotoxin; carcinoma; breast cancer.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9700271-A1.

XX PD 03-JAN-1997.

XX PF 13-JUN-1996; 96WO-US10287.

XX PR 15-JUN-1995; 95US-0000250.

XX PR 14-JUN-1995; 95US-0000238.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Marks JD, Schler R;

XX DR WPI: 1997-077488/07.

XX PS Claim 10; Page 71; 117pp; English.

The present sequence represents a light chain shuffled mutant C6VLB, modified from the derived light chain variable region of the C6 human antibody C6.5. The mutant C6VLB was selected on polystyrene adsorbed antigen, and it binds specifically to c-erbB-2. A chimeric molecule that binds specifically to tumour cells carrying c-erbB-2 consists of an effector compound attached to a C6 human antibody, e.g. C6.5. If the effector compound is a cytotoxin the chimeric molecule can be used to inhibit growth of c-erbB-2 positive tumours (especially breast and other carcinomas). If the effector compound is a label the chimeric molecule can be used to detect such cells, including in vivo localisation. The antibody can also be used for diagnosis/localisation, in vivo or in vitro, especially by immunoassay. The nucleic acid encoding the antibody, and a nucleic acid encoding a single chain polypeptide with the binding specificity of the antibody and comprising the binding portions of variable regions of light and heavy chains of the antibody, joined by a linker, can be used to produce recombinant proteins by standard methods. Unlike known anti-c-erbB-2 antibodies, C6 antibodies are fully human, so should elicit little, if any, immunogenic response.

XX SQ Sequence 112 AA;

Query Match 86.8%; Score 33; DB 18; Length 112;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 DNORPS 7
 111111

Db 52 dngrps 57

RESULT 10

AAAR54051
 ID AAR54051 standard; Protein; 108 AA.

XX AC AAR54051;

XX DT 08-NOV-1994 (first entry)

XX DE Sequence of the VL region of monoclonal antibody PE1-1 against hepatitis B virus surface antigen.

XX DE Hepatitis B virus; surface antigen; monoclonal antibody; therapy;

XX KW HBSAG; diagnosis; HBV.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Region 1..96

FT Region /Label= VL V

FT Region 24..337

FT Region /Label= CDR 1

FT Region 50..56

FT Region /Label= CDR 2

FT Region 90..98

FT Region /Label= CDR 3

FT Region 97..108

FT Region /Label= JL 3

XX PN WO9411495-A.

XX PD 26-MAY-1994.

XX PF 06-NOV-1992; 92WO-US09749.

XX PR 06-NOV-1992; 92WO-US09749.

XX PA (SANO) SANDOZ LTD.

XX PI Ostberg LG;

XX DR WPI: 1994-183497/22.

XX DR N-PSDB: AAO64054.

XX PS Monoclonal antibodies active against Hepatitis B surface antigen - for diagnosis and treatment of Hepatitis B virus

XX Example; Page 40; 53pp; English.

Human volunteers were immunised with hepatitis B vaccine. MD3-4, ZM1-2, ZM1-1 and PE1-1 hybridoma cell lines were derived from lymphocytes of individuals immunised with Hepavax (Merck & Co). Antibodies PE1-1, ZM1-1, ZM1-2 and MD3-4 belong to the IgG1 class. The cell lines producing PE1-1, ZM1-1 and ZM1-2 were deposited as ATCC HB9234, 9191 and 9192 respectively. The cell lines all behave as typical (mouse x human) x human hybridomas and produce their respective Abs in concs. ranging up to 25 mg/l in standard suspension culture. The heavy variable (VH) and light variable (LH) chains of Abs PE1-1, ZM1-1, ZM1-2 and MD3-4 were isolated and sequenced. Total RNA was extracted from 10(7) hybridoma cells of each cell line. ss DNA was synthesised using AMV-reverse transcriptase and oligo-dT as primer. PCR was performed and amplified DNA was size selected. ss DNA for sequencing was isolated from each positive clone after superinfection with M13K07. Sequencing was by the dideoxy chain termination method (Sanger et al.).

XX SQ Sequence 108 AA;

Query Match 84.2%; Score 32; DB 15; Length 108;

Best Local Similarity 71.4%; Pred. No. 21;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
:|||||
DB 49 ddnrps 55

RESULT 11
AAW01523

ID AAW01523 standard; Protein: 108 AA.

AC AAW01523;

DT 04-MAR-1997 (first entry)

DE Monoclonal antibody PE1-1 light chain variable region.

KW Monoclonal antibody; diagnosis; treatment; infection; hepatitis B;
KW xenogeneic hybridoma; SPAZ 4; PE1-1; ZM1-1; ZM1-2; MD3-4; L03-3;
KW IgG1 class; heavy chain; light chain; variable region.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Region 1..96 /label- VV_region

FT Region 24..30 /label- CDR1

FT Region 31..49 /label- Framework_region

FT Region 50..54 /label- CDR2

FT Region 55..89 /label- Framework_region

FT Region 90..98 /label- CDR3

FT Region 97..108 /label- J3_region

PN US556354-A.

PD 15-OCT-1996.

PE 05-SEP-1986; 86US-0904517.

PR 21-APR-1992; 92US-0871426.

PR 05-SEP-1986; 86US-0904517.

PR 31-OCT-1986; 86US-0925196.

PR 11-MAY-1988; 88US-0192754.

PR 15-JUN-1990; 90US-0538796.

PR 27-MAR-1991; 91US-0676036.

PR 14-JUN-1994; 94US-0259372.

PA (SANO) SANDOZ LTD.

PI Ostberg LG;

DR WPI: 1996-476304/47.

DR N-PSDB; AAT46129.

PT Human monoclonal antibodies specific for hepatitis B surface antigen

PT - are used to treat or prevent infection or in diagnostic assays

PS Claim 4: Column 37-38; 26pp; English.

CC Monoclonal antibodies effective for the diagnosis and treatment of
CC diseases caused by infection with hepatitis B have been prepared from a
CC cell line obtained by fusing a xenogeneic hybridoma designated SPAZ 4
CC with blood cells of a patient immunised with hepatitis B vaccine.
CC Specific antibodies are PE1-1, ZM1-1, ZM1-2, MD3-4 and L03-3, each of
CC these being of the IgG1 class. The present sequence is the light
CC variable chain of PE1-1.

XX SQ Sequence 108 AA;

Query Match 84.2%; Score 32; DB 17; Length 108;
Best Local Similarity 71.4%; Pred. No. 21;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
:|||||
DB 49 ddnrps 55

RESULT 12
AAW24988

ID AAW24988 standard; Protein: 108 AA.

AC AAW24988;

DT 08-DEC-1997 (first entry)

DE Monoclonal antibody PE1-1 VI region.

KW Heavy chain; light chain; variable region; human; monoclonal antibody;
KW immunisation; hepatitis B virus; HBV; vaccine; mouse; fusion; xenogeneic;
KW peripheral blood lymphocyte; surface antigen; cell culture; ion exchange;
KW chromatography; size separation; primer; PCR; polymerase chain reaction;
KW amplification; hybridoma; infection; immunosuppression; hepatitis;
KW liver transplant.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Region 1..96 /note- "VI region"

FT Domain 24..33 /note- "complementarity determining region 1"

FT Domain 50..56 /note- "complementarity determining region 2"

FT Domain 90..98 /note- "complementarity determining region 3"

FT Region 97..108 /note- "J13 region"

PN US5648077-A.

PD 15-JUL-1997.

PE 05-SEP-1986; 86US-0904517.

PR 21-APR-1992; 92US-0871426.

PR 05-SEP-1986; 86US-0904517.

PR 31-OCT-1986; 86US-0925196.

PR 11-MAY-1988; 88US-0192754.

PR 15-JUN-1990; 90US-0538796.

PR 27-MAR-1991; 91US-0676036.

PR 14-JUN-1994; 94US-0259372.

PR 06-JUN-1995; 95US-0468671.

PA (SANO) SANDOZ LTD.

PI Ostberg LG;

DR WPI: 1997-372021/34.

DR N-PSDB; AAT85842.

PT Treatment of hepatitis B - with human monoclonal antibody

PS Example 8: Column 19-22; 25pp; English.

CC This is the amino acid sequence of the light chain variable (VI) region
CC from the human monoclonal antibody (Mab) PE1-1. The Mab was generated
CC by immunising humans with a hepatitis B virus (HBV) vaccine, isolating

peripheral blood lymphocytes (PBL) and fusing them with a mouse/human xenogeneic cell line SP2-4. 5 cell lines were isolated: Pel-1, ZM1-1, ZM1-2, MD3-4 and IO3-3. The cell lines were tested for production of an anti-hepatitis B virus surface antigen antibody by ELISA. The Mabs are then purified from large scale cell culture by protein A chromatography, size separation on Sephacryl S500 gel and ion exchange chromatography on Q-Sepharose. The heavy and light chains of the Mabs were isolated and their amino acid sequences determined. Primers were generated and used to amplify cDNA synthesised from RNA purified from each hybridoma cell line. The sequences of the heavy and light chains (nucleic acid and amino acid) from Mabs Pel-1, ZM1-1, ZM1-2 and MD3-4 are shown in AA85838-45 and AA824984-91. The Mabs can be used to treat HBV infections in immunosuppressed patients or patients with chronic active hepatitis, especially liver transplant patients.

SQ Sequence 108 AA;

Query Match	84.2%;	Score 32;	DB 18;	Length 108;
Best local Similarity	71.4%;	Pred. No. 21;		
Matches	5;	Conservative	2;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy	1	EDNQPS	7
	:	:	
Db	49	ddnerps	55

RESULT	13
AAR88719	
ID	AAR88719 standard; protein; 108 AA

AC AAR88719;

DT 26-SEP-1996 (first entry)

DE Human antibody lambda chain consensus sequence.

KW Antibody; stability; instability; mutagenesis; amino acid frequency;

OS Homo sapiens.

PN DE4425115-A1.

PD 18-JAN-1996.

PF 15-JUL-1994; 94DE-4425115.

15-JUL-1994; 94DE-4425115.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Steinhilber S, Steinhilber B:

WP1: 1996-069594/08

PT Functional antibodies modified to increase or decrease stability -
PT have specific amino acid substns. in the variable domain, determined
PT by reference to consensus sequences

PS Claim 2; Page 31; 43pp; German.

Consensus sequences were established for antibody variable regions from human and mouse heavy and light (κ and λ) chains. In addition, frequency tables, listing the amino acids which can be found at each position were compiled. In a new method, mutations are introduced into an antibody variable domain by reference to the frequency table. When an amino acid is replaced by one which occurs at a higher frequency at that position, the resulting antibody chain is more stable than the wild-type; when an amino acid is replaced by one which occurs less frequently (or not at all), the resulting

CC antibody is less stable than the wild-type. Stabilised antibodies
CC are useful as diagnostic reagents, as catalysts and in treatment of
CC cancer, autoimmune diseases and infections. Destabilised antibodies
CC have improved (faster) pharmacokinetic properties.
CC The present sequence is the human lambda light chain variable region
CC consensus sequence.

SQ Sequence 108 AA;

Query Match	81.6%;	Score 31;	DB 17;	Length 108;
Best Local Similarity	71.4%;	Pred. No. 34;		
Matches	5;	Conservative	2;	Mismatches 0;
				Indels 0;
				Gaps 0.

QY	1	EDNQPS	7
		:11:111	
Db	49	ddnkrrps	55

RESULT	14
AAV68898	
ID	AAV68898 standard; Protein; 108 AA

AC AAY68898;

DT 16-MAY-2000 (first entry)

DE The variable region of the light chain of HMab503.

KW variable light chain; human anti-HCV antibody; chronic Hepatitis C;

KW liver transpi

... Homo sapiens.

PN W0200005266-

03 - FEB - 2000.

20-JUL-1999: 99WO-EP05173

21-JUL-1998: 98EP-0113595.

XX
PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG &

PA (JNRM) INSERM INST NAI SANIE & RECH MEDICALE.
XY

PI	Reiter C
PI	Tschannen

XX WPT, 2000-192654/16
XX DP

DR N-PsDB; AA260759

AA	PT	PT	PT
Novel antibodies, antigens, for diagnosing and treating Hepatitis C	Virus, comprising at least one complementarity determining region of	the variable domain of a human antibody -	

PS Claim 3; Fig 5; 64pp; English

The present sequence represents the variable light chain of a human anti-Hepatitis C virus (HCV) antibody. The sequence was used to construct anti-HCV antibodies of the invention. These antibodies comprise at least one complementarily determining region (CDR) of the variable domain of a human antibody that specifically recognizes a conformation-dependent epitope of HCV glycoprotein E2 and is capable of precipitating covalently or non-covalently associated E2/E1 complexes. The antibodies are useful for preventing (re)infection of HCV and are useful for alleviating chronic Hepatitis C in a human or an animal. The antibodies are useful for diagnosing chronic Hepatitis C for the presence of neutralization of binding of HCV glycoprotein E2 to target cells. The antibodies are useful for the treatment or prevention of HCV infection or recurrence of HCV infection, in which they are administered prior, during or after liver transplantation. Therapeutic compositions containing the antibodies are also useful.

CC in liver transplantation. The antibody is also useful for the prevention
 CC of infection of Tupaia-hepatocyte with HCV-infectious human sera.
 XX
 SQ Sequence 108 AA;

Query Match 81.6%; Score 31; DB 21; Length 108;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EDNRPS 7
 :||:||||
 Db 49 kdnerps 55

RESULT 15
 AAY68900
 ID AAY68900 standard; Protein: 108 AA.
 XX
 AC AAY68900;

DT 16-MAY-2000 (first entry)
 XX

DE Variable region of the light chain of an anti-HCV antibody.
 XX

KM Variable light chain; human anti-HCV antibody; chronic Hepatitis C;
 KM conformation-dependent epitope; HCV glycoprotein E2; HCV infection;
 KM liver transplantation; Tupaia-hepatocyte; HCV-infectious human sera.
 XX

OS Homo sapiens.
 XX

PN W0200005266-A1.
 XX

PD 03-FEB-2000.
 XX

PF 20-JUL-1999; 99WC-EP05173.
 XX

PR 21-JUL-1998; 98EP-0113595.
 XX

PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.
 PA (INRM) INSERM NAT SANTE & RECH MEDICALE.
 XX

PI Reiter C, Habersetzer F, Fournillier A, Trepo C, Desgranges C;
 PI Inchauspe G;
 XX

DR WPI; 2000-182654/16.
 DR N-PSDB; AA260761.
 XX

PT Novel antibodies, antigens, for diagnosing and treating Hepatitis C
 PT virus, comprising at least one complementarily determining region of
 PT the variable domain of a human antibody -
 XX

PS Claim 3; Page 58; 64pp; English.
 XX

CC The present sequence represents the variable light chain of a human
 CC anti-Hepatitis C virus (HCV) antibody. The sequence was used to
 CC construct antibodies of the invention. These antibodies comprise at
 CC least one complementarity determining region (CDR) of the variable
 CC domain of a human antibody that specifically recognizes a
 CC conformation-dependent epitope of HCV glycoprotein E2 and is capable
 CC of precipitating covalently or non-covalently associated E2/E1
 CC complexes. The antibodies are useful for preventing (re)infection of
 CC HCV and are useful for alleviating chronic Hepatitis (re)infection of
 CC an animal. The antibodies are useful for diagnosing chronic Hepatitis C
 CC to the presence of neutralization of binding of HCV glycoprotein E2 on
 CC to tablet cells. The antibodies are useful for the treatment or
 CC prevention of HCV infection or reoccurrence of HCV infection, in which
 CC they are administered prior, during or after liver transplantation.
 CC Therapeutic compositions containing the antibodies are also useful
 CC in liver transplantation. The antibody is also useful for the prevention
 CC of infection of Tupaia-hepatocyte with HCV-infectious human sera.
 XX

SQ Sequence 108 AA;

Query Match 81.6%; Score 31; DB 21; Length 108;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EDNRPS 7
 :||:||||
 Db 49 kdnerps 55

Search completed: November 19, 2001, 08:12:36
 Job time: 89 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:16 ; Search time 526.89 seconds
(without alignments)
3.689 Million cell updates/sec

Title: US-09-610-118-65

Perfect score: 38

Sequence: 1 EDNORPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	1	PCR-US00-18152-65
2	38	100.0	7	20	US-09-610-118-65
3	38	100.0	7	22	US-09-832-312-65
4	38	100.0	79	24	US-60-163-062-790
5	38	100.0	79	24	US-60-163-062-948
6	38	100.0	88	24	US-60-195-053-2358
7	38	100.0	94	24	US-60-163-123-1181
8	38	100.0	94	24	US-60-163-123-1427
9	38	100.0	94	24	US-60-195-053-2359

10	38	100.0	94	24	US-60-234-446-712	Sequence 712, App
11	38	100.0	95	24	US-60-162-243-594	Sequence 594, App
12	38	100.0	95	24	US-60-162-247-4885	Sequence 4885, App
13	38	100.0	100	24	US-60-196-710-4905	Sequence 4905, App
14	38	100.0	100	24	US-60-196-718-4971	Sequence 4971, App
15	38	100.0	104	24	US-60-162-247-3052	Sequence 3052, App
16	38	100.0	104	24	US-60-162-247-3489	Sequence 3489, App
17	38	100.0	108	24	US-60-169-840-7123	Sequence 7123, App
18	38	100.0	108	24	US-60-169-840-7362	Sequence 7362, App
19	38	100.0	109	24	US-60-169-868-5171	Sequence 5171, App
20	38	100.0	111	24	US-60-162-243-421	Sequence 421, App
21	38	100.0	111	24	US-60-162-247-4435	Sequence 4435, App
22	38	100.0	111	24	US-60-169-840-8023	Sequence 8023, App
23	38	100.0	111	24	US-60-169-868-5926	Sequence 5926, App
24	38	100.0	161	1	PCT-US00-265248-6775	Sequence 6775, App
25	38	100.0	218	21	US-09-760-479-543	Sequence 543, App
26	38	100.0	258	1	PCT-US01-19110-1234	Sequence 1234, App
27	38	100.0	258	22	US-09-880-748-1234	Sequence 1234, App
28	34	89.5	7	14	US-09-043-522A-13	Sequence 13, App
29	34	89.5	7	14	US-09-043-530-13	Sequence 13, App
30	34	89.5	101	1	PCT-US01-10472-85	Sequence 85, App
31	34	89.5	109	14	US-09-043-522A-9	Sequence 9, App
32	34	89.5	109	14	US-09-043-530-9	Sequence 9, App
33	34	89.5	110	14	US-09-043-522A-24	Sequence 24, App
34	34	89.5	110	14	US-09-043-530-24	Sequence 24, App
35	34	89.5	114	1	PCT-US00-05883-1194	Sequence 1194, App
36	34	89.5	114	23	US-09-925-299-1194	Sequence 1194, App
37	34	89.5	117	21	US-09-760-476-1829	Sequence 1829, App
38	34	89.5	117	21	US-09-760-479-771	Sequence 771, App
39	34	89.5	233	24	US-60-222-372-1	Sequence 1, App
40	34	89.5	244	1	PCT-US01-19110-1842	Sequence 1842, App
41	34	89.5	244	22	US-09-880-748-1842	Sequence 1842, App
42	34	89.5	246	1	PCT-US01-10472-252	Sequence 252, App
43	34	89.5	112	17	US-09-315-574-39	Sequence 39, App
44	33	86.8	248	1	PCT-US01-19110-1273	Sequence 1273, App
45	33	86.8	248	22	US-09-880-748-1273	Sequence 1273, App

ALIGNMENTS

RESULT 1

Sequence 65, Application PC/TUS0018152

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.

TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF

FILE REFERENCE: 7853-211-228

CURRENT APPLICATION NUMBER: PCT/US00/18152

CURRENT FILING DATE: 2000-06-30

EARLIER APPLICATION NUMBER: 09/503,387

EARLIER FILING DATE: 2/14/00

EARLIER APPLICATION NUMBER: 09/454,824

EARLIER FILING DATE: 12/6/99

EARLIER APPLICATION NUMBER: 09/345,468

NUMBER OF SEQ ID NOS: 72

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 65

LENGTH: 7

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US00-18152-65

Query Match 100.0% Score 38; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
|||||
DB 1 EDNORPS 7

```
RESULT 2
US-09-610-118-65
; Sequence 65, Application US/09610118
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villaval, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Valinchenker, W.
; APPLICANT: Gill, D.
; APPLICANT: Qian, M.
; APPLICANT: Kingsbury, G.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-211
; CURRENT APPLICATION NUMBER: US/09/610,118
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2/14/00
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 12/6/99
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 6/30/99
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 65
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-610-118-65

Query Match          100.0%; Score 38; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
Db 1 EDNORPS 7

RESULT 3
US-09-832-312-65
; Sequence 65, Application US/09832312
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 65
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-65

Query Match          100.0%; Score 38; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
Db 1 EDNORPS 7
```

```
RESULT 4
US-60-163-062-790
; Sequence 790, Application US/60163062
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CLO00134
; CURRENT APPLICATION NUMBER: US/60/163,062
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1302
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 790
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Human
US-60-163-062-790

Query Match          100.0%; Score 38; DB 24; Length 79;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
Db 37 EDNORPS 43

RESULT 5
US-60-163-062-948
; Sequence 948, Application US/60163062
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CLO00134
; CURRENT APPLICATION NUMBER: US/60/163,062
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1302
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 948
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Human
US-60-163-062-948

Query Match          100.0%; Score 38; DB 24; Length 79;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
Db 37 EDNORPS 43

RESULT 6
US-60-195-053-2358
; Sequence 2358, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CLO00427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2358
; LENGTH: 88
; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-2358
```

```
Query Match
Best Local Similarity 100.0%; Score 38; DB 24; Length 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EDNRPS 7
      |||||
Db 73 EDNRPS 79
```

```
RESULT 7
US-60-163-123-1181
; Sequence 1181, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000137
; CURRENT APPLICATION NUMBER: US/60/163,123
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1181
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Human
US-60-163-123-1181
```

```
Query Match
Best Local Similarity 100.0%; Score 38; DB 24; Length 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EDNRPS 7
      |||||
Db 53 EDNRPS 59
```

```
RESULT 8
US-60-163-123-1427
; Sequence 1427, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000137
; CURRENT APPLICATION NUMBER: US/60/163,123
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1427
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Human
US-60-163-123-1427
```

```
Query Match
Best Local Similarity 100.0%; Score 38; DB 24; Length 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EDNRPS 7
      |||||
Db 53 EDNRPS 59
```

```
RESULT 9
US-60-195-053-2359
; Sequence 2359, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: C1000427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2359
; LENGTH: 94
; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-2359
```

```
Query Match
Best Local Similarity 100.0%; Score 38; DB 24; Length 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EDNRPS 7
      |||||
Db 70 EDNRPS 76
```

```
RESULT 10
US-60-234-446-712
; Sequence 712, Application US/60234446
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: C1000832
; CURRENT APPLICATION NUMBER: US/60/234,446
; CURRENT FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 1797
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 712
; LENGTH: 94
; TYPE: PRT
; ORGANISM: HUMAN
US-60-234-446-712
```

```
Query Match
Best Local Similarity 100.0%; Score 38; DB 24; Length 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EDNRPS 7
      |||||
Db 53 EDNRPS 59
```

```
RESULT 11
US-60-162-243-594
; Sequence 594, Application US/60162243
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000132
; CURRENT APPLICATION NUMBER: US/60/162,243
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 828
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
```

; LENGTH: 95
; TYPE: PRT
; ORGANISM: Human
US-60-162-243-594

Query Match 100.0%; Score 38; DB 24; Length 95;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
Db 39 EDNORPS 45

RESULT 12
US-60-162-247-4885
; Sequence 4885, Application US/60162247
; GENERAL INFORMATION:
; APPLICANT: Bonazzl, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CLO00127
; CURRENT APPLICATION NUMBER: US/60/162,247
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 5442
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4885
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Human
US-60-162-247-4885

Query Match 100.0%; Score 38; DB 24; Length 95;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
Db 39 EDNORPS 45

RESULT 13
US-60-196-710-4905
; Sequence 4905, Application US/60196710
; GENERAL INFORMATION:
; APPLICANT: Bonazzl, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: CLO00450
; CURRENT APPLICATION NUMBER: US/60/196,710
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7166
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4905
; LENGTH: 100
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-710-4905

Query Match 100.0%; Score 38; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
Db 73 EDNORPS 79

RESULT 14
US-60-196-718-4971
; Sequence 4971, Application US/60196718
; GENERAL INFORMATION:
; APPLICANT: Bonazzl, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CLO00456
; CURRENT APPLICATION NUMBER: US/60/196,718
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4971
; LENGTH: 100
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-718-4971

Query Match 100.0%; Score 38; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
Db 73 EDNORPS 79

RESULT 15
US-60-162-247-3052
; Sequence 3052, Application US/60162247
; GENERAL INFORMATION:
; APPLICANT: Bonazzl, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CLO00127
; CURRENT APPLICATION NUMBER: US/60/162,247
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 5442
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3052
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Human
US-60-162-247-3052

Query Match 100.0%; Score 38; DB 24; Length 104;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
Db 51 EDNORPS 57

Search completed: November 19, 2001, 08:23:16
Job time: 729 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: November 19, 2001, 08:23:32 ; Search time 11.44 Seconds

(without alignments)
3.259 Million cell updates/sec

Title: US-09-610-118-65

Perfect score: 38

Sequence: 1 EDNORPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 25229 seqs, 5326477 residues

Total number of hits satisfying chosen parameters: 25229

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents, AA, New: *
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep: *
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep: *
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep: *
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep: *
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep: *
6: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	5	US-09-972-656-13
2	38	100.0	103	5	US-09-972-656-130
3	38	100.0	216	5	US-09-972-656-108
4	33	86.8	7	5	US-09-972-656-14
5	33	86.8	7	5	US-09-972-656-20
6	33	86.8	216	5	US-09-972-656-90
7	33	86.8	218	5	US-09-972-656-102
8	29	76.3	7	5	US-09-972-656-12
9	29	76.3	217	5	US-09-972-656-88
10	28	73.7	176	5	US-09-972-656-5795
11	28	73.7	305	5	US-09-972-656-12466
12	28	73.7	347	5	US-09-972-656-12466
13	28	73.7	1317	5	US-09-972-656-12466
14	27	71.1	246	5	US-09-972-656-7544
15	27	71.1	314	5	US-09-972-656-11924
16	27	71.1	366	5	US-09-972-656-7434
17	27	71.1	508	5	US-09-972-656-7434
18	27	71.1	520	5	US-09-972-656-7434
19	27	71.1	1041	5	US-09-972-656-7434
20	27	71.1	1041	5	US-09-972-656-7434
21	27	71.1	1041	5	US-09-972-656-7434
22	27	71.1	1041	5	US-09-972-656-7434
23	26	68.4	26	5	US-09-972-656-7434
24	26	68.4	35	5	US-09-972-656-7434
25	26	68.4	171	5	US-09-972-656-7434
26	26	68.4	304	5	US-09-972-656-7434
27	26	68.4	331	5	US-09-972-656-7434

28	26	68.4	496	5	US-09-815-242-12203	Sequence 12203, A
29	26	68.4	744	5	US-09-815-242-13390	Sequence 13390, A
30	26	68.4	850	1	PCT-US01-07265-4	Sequence 4, App1
31	26	68.4	850	5	US-09-799-629-4	Sequence 4, App1
32	26	68.4	904	5	US-09-964-899-49	Sequence 49, App1
33	26	68.4	904	5	US-09-976-594-324	Sequence 324, App
34	26	68.4	984	5	US-09-976-594-127	Sequence 127, App
35	25	65.8	7	5	US-09-972-656-17	Sequence 17, App1
36	25	65.8	148	5	US-09-972-656-18	Sequence 18, App1
37	25	65.8	153	5	US-09-897-516-8284	Sequence 8284, App
38	25	65.8	160	5	US-09-897-516-6009	Sequence 6009, App
39	25	65.8	212	5	US-09-972-656-98	Sequence 98, App1
40	25	65.8	214	5	US-09-972-656-96	Sequence 96, App1
41	25	65.8	234	5	US-09-576-424-2	Sequence 2, App1
42	25	65.8	268	1	PCT-US01-31857-1	Sequence 1, App1
43	25	65.8	268	5	US-09-976-118-1	Sequence 1, App1
44	25	65.8	300	5	US-09-897-516-4820	Sequence 4820, App
45	25	65.8	300	5	US-09-897-516-4820	Sequence 4820, App

ALIGNMENTS

```
RESULT 1
US-09-972-656-13
; Sequence 13, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-13

Query Match      100.0%; Score 38; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDNORPS 7
Db      1 EDNORPS 7

RESULT 2
US-09-972-656-130
; Sequence 130, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Misc.
; LOCATION: (7)..(7)
; OTHER INFORMATION: Unidentified
```

```

; NAME/KEY: Misc.
; LOCATION: (33)..(33)
; OTHER INFORMATION: Unidentified
; NAME/KEY: Misc.
; LOCATION: (34)..(34)
; OTHER INFORMATION: Unidentified
; NAME/KEY: Misc.
; LOCATION: (35)..(35)
; OTHER INFORMATION: Unidentified
; NAME/KEY: Misc.
; LOCATION: (36)..(36)
; OTHER INFORMATION: Unidentified
US-09-972-656-130
```

```
Query Match          100.0%; Score 38; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 EDDQRPS 7
    |||||||
Db 56 EDDQRPS 62
```

```
RESULT 3
US-09-972-656-108
; Sequence 108, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-108
```

```
Query Match          100.0%; Score 38; DB 5; Length 216;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 EDDQRPS 7
    |||||||
Db 51 EDDQRPS 57
```

```
RESULT 4
US-09-972-656-14
; Sequence 14, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-14
```

```
Query Match          86.8%; Score 33; DB 5; Length 7;
Best Local Similarity 85.7%; Pred. No. 2.2e+04;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 EDDQRPS 7
    |||||||
Db 1 EDDQRPS 7
```

```
RESULT 5
US-09-972-656-20
; Sequence 20, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-20
```

```
Query Match          86.8%; Score 33; DB 5; Length 7;
Best Local Similarity 85.7%; Pred. No. 2.2e+04;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 EDDQRPS 7
    |||||||
Db 1 EDDQRPS 7
```

```
RESULT 6
US-09-972-656-90
; Sequence 90, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-90
```

```
Query Match          86.8%; Score 33; DB 5; Length 216;
Best Local Similarity 85.7%; Pred. No. 1.1;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 EDDQRPS 7
    |||||||
Db 51 EDDQRPS 57
```

```
RESULT 7
US-09-972-656-102
; Sequence 102, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
```

```

; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 102
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-102
```

```

Query Match
Best Local Similarity 86.8%; Score 33; DB 5; Length 218;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORPS 7
Db 51 EDDORPS 57
```

```

RESULT 8
US-09-972-656-12
; Sequence 12, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-12
```

```

Query Match
Best Local Similarity 76.3%; Score 29; DB 5; Length 7;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 EDNORPS 7
Db 1 EDKERPS 7
```

```

RESULT 9
US-09-972-656-88
; Sequence 88, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; APPLICANT: Tsai, Mei-Mei
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972,656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 88
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-88
```

```

Query Match 76.3%; Score 29; DB 5; Length 217;
```

```

Best Local Similarity 71.4%; Pred. No. 7.6;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 EDNORPS 7
Db 51 EDKERPS 57
```

```

RESULT 10
US-09-897-516-5795
; Sequence 5795, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karlina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 5795
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-5795
```

```

Query Match
Best Local Similarity 73.7%; Score 28; DB 5; Length 176;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 EDNORPS 7
Db 116 ENSORPS 122
```

```

RESULT 11
US-09-815-242-5706
; Sequence 5706, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
```

```

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5706
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5706
```

```

Query Match          73.7%; Score 28; DB 5; Length 305;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 EDNORP 6
        |||:|
Db      15 EDNEKP 20
```

```

RESULT 12
US-09-815-242-12466
; Sequence 12466, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12466
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12466
```

```

Query Match          73.7%; Score 28; DB 5; Length 347;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 EDNORP 6
        |||:|
Db      15 EDNEKP 20
```

```

RESULT 13
US-09-815-242-5118
; Sequence 5118, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
```

```

; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5118
; LENGTH: 1317
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5118
```

```

Query Match          73.7%; Score 28; DB 5; Length 1317;
Best Local Similarity 71.4%; Pred. No. 81;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 EDNORPS 7
        |||:|
Db      831 DDNORPS 837
```

```

RESULT 14
US-09-897-516-7544
; Sequence 7544, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 7544
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-7544
```

```

Query Match          71.1%; Score 27; DB 5; Length 246;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

OY 2 DNQRP 7
111 11
Db 233 DNQRP 238

RESULT 15

US-09-815-242-11924
; Sequence 11924, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11924
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11924

Query Match 71.1%; Score 27; DB 5; Length 314;
Best Local Similarity: 83.3%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 1; Indels 0;

OY 1 EDNORP 6
1111 1
Db 63 EDNORP 68

Search completed: November 19, 2001, 08:23:32
Job time: 744 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:13:25 ; Search time 43.5 Seconds
(without alignments)
3.621 Million cell updates/sec

Title: US-09-610-118-65

Perfect score: 38

Sequence: 1 EDNORPS 7

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	89.5	106	1 US-08-259-372A-16	Sequence 16, App1
2	34	89.5	106	1 US-08-468-671-16	Sequence 16, App1
3	33	86.8	112	2 US-08-665-202-19	Sequence 39, App1
4	32	84.2	108	1 US-08-259-372A-10	Sequence 10, App1
5	32	84.2	108	1 US-08-468-671-10	Sequence 10, App1
6	31	81.6	64	2 US-08-765-179B-19	Sequence 19, App1
7	31	81.6	108	4 US-09-025-769B-20	Sequence 20, App1
8	31	81.6	109	4 US-09-157-370-5	Sequence 5, App1
9	30	78.9	109	4 US-09-025-769B-32	Sequence 32, App1
10	30	78.9	109	4 US-09-025-769B-51	Sequence 51, App1
11	30	78.9	462	3 US-08-875-944B-5	Sequence 4, App1
12	30	78.9	462	2 US-09-116-049-4	Sequence 4, App1
13	30	78.9	555	4 US-08-780-835B-2	Sequence 2, App1
14	30	78.9	555	4 US-09-303-268-2	Sequence 2, App1
15	30	78.9	555	4 US-09-116-049-2	Sequence 2, App1
16	30	78.9	620	5 PCT-US95-05008-9	Sequence 9, App1
17	29	76.3	109	1 US-08-478-039-91	Sequence 91, App1
18	29	76.3	109	1 US-08-478-039-91	Sequence 91, App1
19	29	76.3	111	4 US-08-983-607-31	Sequence 31, App1
20	29	76.3	112	4 US-08-983-607-31	Sequence 31, App1
21	29	76.3	112	4 US-08-983-607-31	Sequence 31, App1
22	29	76.3	132	2 US-08-534-975-4	Sequence 4, App1
23	29	76.3	132	2 US-08-534-975-4	Sequence 4, App1
24	29	76.3	132	2 US-09-129-855A-4	Sequence 4, App1
25	28	73.7	7	1 US-09-247-154-4	Sequence 4, App1
26	28	73.7	7	1 US-08-264-093-19	Sequence 19, App1
27	28	73.7	98	1 US-08-211-202-111	Sequence 111, App1
28	28	73.7	98	2 US-08-665-202-38	Sequence 38, App1

28	28	73.7	109	4	US-09-240-274-54	Sequence 54, App1
29	28	73.7	109	4	US-09-240-274-58	Sequence 58, App1
30	28	73.7	109	4	US-09-240-274-59	Sequence 59, App1
31	28	73.7	109	4	US-09-240-274-180	Sequence 180, App1
32	28	73.7	110	4	US-09-240-274-60	Sequence 60, App1
33	28	73.7	110	4	US-09-240-274-63	Sequence 63, App1
34	28	73.7	111	2	US-08-264-093-6	Sequence 6, App1
35	28	73.7	111	2	US-08-665-202-41	Sequence 41, App1
36	28	73.7	111	4	US-08-983-607-21	Sequence 21, App1
37	28	73.7	111	4	US-08-983-607-23	Sequence 23, App1
38	28	73.7	113	1	US-08-211-202-112	Sequence 112, App1
39	28	73.7	113	1	US-08-211-202-113	Sequence 113, App1
40	28	73.7	113	1	US-08-211-202-114	Sequence 114, App1
41	28	73.7	131	1	US-08-305-683A-4	Sequence 4, App1
42	28	73.7	1367	1	US-07-813-593-4	Sequence 4, App1
43	28	73.7	1367	1	US-07-977-451-6	Sequence 6, App1
44	28	73.7	1367	1	US-07-946-507-4	Sequence 4, App1
45	28	73.7	1367	1	US-08-252-517-6	Sequence 6, App1

ALIGNMENTS

RESULT 1
US-08-259-372A-16
Sequence 16, Application US/08259372A
Patent No. 5565354
GENERAL INFORMATION:
APPLICANT: Ostberg, Lars G.
TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,372A
FILING DATE: 14-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/871,426
FILING DATE: 21-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,036
FILING DATE: 27-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/538,796
FILING DATE: 15-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/192,754
FILING DATE: 11-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/925,196
FILING DATE: 31-OCT-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/904,517
FILING DATE: 05-SEP-1986
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-50-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 576-0300

;; INFORMATION FOR SEQ ID NO: 16:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 106 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-259-372A-16

Query Match 89.5%; Score 34; DB 1; Length 106;
Best Local Similarity 85.7%; Pred. No. 4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
1111111
DB 49 EDNRPS 55

RESULT 2
US-08-468-671-16
; Sequence 16, Application US/08468671
; Patent No. 5648077
; GENERAL INFORMATION:
; APPLICANT: Osberg, Lars G.
; TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,671
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/259,372
; FILING DATE: 14-JUN-1994
; APPLICATION NUMBER: US 07/871,426
; FILING DATE: 21-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,036
; FILING DATE: 27-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/538,796
; FILING DATE: 15-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/192,754
; FILING DATE: 11-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/925,196
; FILING DATE: 31-OCT-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/904,517
; FILING DATE: 05-SEP-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-50-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids

;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-468-671-16

Query Match 89.5%; Score 34; DB 1; Length 106;
Best Local Similarity 85.7%; Pred. No. 4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNRPS 7
1111111
DB 49 EDNRPS 55

RESULT 3
US-08-665-202-39
; Sequence 39, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-665-202-39

Query Match 86.8%; Score 33; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DNQRPS 7
1111111
DB 52 DNQRPS 57

RESULT 4
US-08-259-372A-10
; Sequence 10, Application US/08259372A
; Patent No. 5565354
; GENERAL INFORMATION:
; APPLICANT: Ostberg, Lars G.
; TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,372A
; FILING DATE: 14-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/871,426
; FILING DATE: 21-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,036
; FILING DATE: 27-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/538,796
; FILING DATE: 15-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/192,754
; FILING DATE: 11-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/925,196
; FILING DATE: 31-OCT-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/904,517
; FILING DATE: 05-SEP-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-259-372A-10

Query Match 84.2%; Score 32; DB 1; Length 108;
Best Local Similarity 71.4%; Pred. No. 10;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
; :||:|
Db 49 DDNERPS 55

RESULT 5
US-08-468-671-10
; Sequence 10, Application US/08468671
; Patent No. 5648077
; GENERAL INFORMATION:

; APPLICANT: Ostberg, Lars G.
; TITLE OF INVENTION: PRODUCTION OF HUMAN MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR HEPATITIS B SURFACE ANTIGEN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,671
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/259,372
; FILING DATE: 14-JUN-1994
; APPLICATION NUMBER: US 07/871,426
; FILING DATE: 21-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,036
; FILING DATE: 27-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/538,796
; FILING DATE: 15-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/192,754
; FILING DATE: 11-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/925,196
; FILING DATE: 31-OCT-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/904,517
; FILING DATE: 05-SEP-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-671-10

Query Match 84.2%; Score 32; DB 1; Length 108;
Best Local Similarity 71.4%; Pred. No. 10;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
; :||:|
Db 49 DDNERPS 55

RESULT 6
US-08-765-179B-19
; Sequence 19, Application US/08765179B
; Patent No. 5854027
; GENERAL INFORMATION:
; APPLICANT: STEINBACHER, Stefan
; TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY

TITLE OF INVENTION: OF ANTIBODIES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Kamei, Murrey & Oram LLP
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,179B
FILING DATE: 14-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02626
FILING DATE: 06-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 44 25 115.7
FILING DATE: 15-JUL-1994
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 64 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-179B-19

Query Match 81.6%; Score 31; DB 2; Length 64;
Best Local Similarity 71.4%; Pred. No. 9.3;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 19 DDKRPS 25

RESULT 7
US-09-025-769B-20
Sequence 20, Application US/09025769B
Patent No. 6300064
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Ilag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
APPLICANT: Plueckhunn, Andreas
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95 11 3021.0

FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-769B-20

Query Match 81.6%; Score 31; DB 4; Length 108;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 48 DDKRPS 54

RESULT 8
US-09-157-370-5
Sequence 5, Application US/09157370A
Patent No. 6262238
GENERAL INFORMATION:
APPLICANT: STEIRE, Boris
APPLICANT: STEINBACHER, Stefan
TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY OF ANTIBODIES
FILE REFERENCE: P8341-8072
CURRENT APPLICATION NUMBER: US/09/157,370A
CURRENT FILING DATE: 1998-09-21
EARLIER APPLICATION NUMBER: 08/765,179
EARLIER FILING DATE: 1997-01-14
EARLIER APPLICATION NUMBER: PCT/EP95/02626
EARLIER FILING DATE: 1995-07-06
EARLIER APPLICATION NUMBER: DE/P44 25 115.7
EARLIER FILING DATE: 1994-07-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
US-09-157-370-5

Query Match 81.6%; Score 31; DB 4; Length 109;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
:|||||
DB 49 DDKRPS 55

RESULT 9
US-09-025-769B-32
Sequence 32, Application US/09025769B
Patent No. 6300064
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Ilag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
APPLICANT: Plueckhunn, Andreas

```

; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-769B-32

Query Match          78.9%  Score 30; DB 4; Length 109;
Best Local Similarity 71.4%  Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 DNNORPS 57

RESULT 10
US-09-025-769B-51
; Sequence 51, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B

```

```

; FILING DATE: 18-FEB-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-769B-51

Query Match          78.9%  Score 30; DB 4; Length 109;
Best Local Similarity 71.4%  Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
DB 51 DNNORPS 57

RESULT 11
US-08-875-944B-5
; Sequence 5, Application US/08875944B
; Patent No. 6096542
; GENERAL INFORMATION:
; APPLICANT: FUJINAGA, Kei
; APPLICANT: YOSHIDA, Koichi
; APPLICANT: HIGASHINO, Fumihito
; TITLE OF INVENTION: CANCER CONTROL
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,944B
; FILING DATE: 07-AUG-1997
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 07-020173
; FILING DATE: 08-FEB-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/00016
; FILING DATE: 09-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: FUJINAGA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-875-944B-5

Query Match
Best Local Similarity 83.3%; Score 30; DB 3; Length 462;
Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
|||||
DB 412 DNORPA 417

RESULT 12
US-09-116-049-4
Sequence 4, Application US/09116049A
Patent No. 6248351
GENERAL INFORMATION:
APPLICANT: Hung, Men-Chle
TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS
FILE REFERENCE: UTSC:562
CURRENT APPLICATION NUMBER: US/09/116,049A
CURRENT FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 462
TYPE: PRT
ORGANISM: Homo sapiens
US-09-116-049-4

Query Match
Best Local Similarity 83.3%; Score 30; DB 4; Length 462;
Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
|||||
DB 412 DNORPA 417

RESULT 13
US-08-780-835B-2
Sequence 2, Application US/08780835B
Patent No. 5922688
GENERAL INFORMATION:
APPLICANT: Hung, Mlen-Chle
APPLICANT: Xing, Xiangming
TITLE OF INVENTION: PEA3 Is a Tumor Suppressor
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780,835B
FILING DATE: 10-JAN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: UTSC500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-780-835B-2

Query Match
Best Local Similarity 83.3%; Score 30; DB 2; Length 555;
Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
|||||
DB 505 DNORPA 510

RESULT 14
US-09-303-268-2
Sequence 2, Application US/09303268
Patent No. 6172212
GENERAL INFORMATION:
APPLICANT: Hung, Mlen-Chle
TITLE OF INVENTION: PEA3 Is a Tumor Suppressor
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,268
FILING DATE: 30-Apr-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/780,835
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: UTSC500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-303-268-2

Query Match
Best Local Similarity 83.3%; Score 30; DB 4; Length 555;
Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
|||||
DB 505 DNORPA 510

```

RESULT 15
US-09-116-049-2
; Sequence 2, Application US/09116049A
; Patent No. 6248351
; GENERAL INFORMATION:
; APPLICANT: Hung, Men-Chie
; TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS
; FILE REFERENCE: UFG:582
; CURRENT APPLICATION NUMBER: US/09/116,049A
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-116-049-2

```

```

Query Match      78.9%; Score 30; DB 4; Length 555;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY      2 DNORPS 7
      |||||
Db      505 DNORPA 510

```

Search completed: November 19, 2001, 08:13:26
 Job time: 139 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:14:22 ; Search time 48.99 Seconds

(without alignments)
10.884 Million cell updates/sec

Title: US-09-610-118-65

Sequence: 1 EDNQRPS 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 segs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	92.1	112	1	L6HUA
2	34	89.5	109	1	L6HUA
3	34	89.5	111	1	L6HUA
4	34	89.5	131	1	L6HUEB
5	34	89.5	266	2	S72947
6	34	89.5	704	2	T02558
7	33	86.8	111	1	L6HUA
8	32	84.2	111	1	L6HUA
9	32	84.2	460	2	T27770
10	32	84.2	471	2	T20690
11	31	81.6	107	2	PC4283
12	31	81.6	161	1	E81447
13	30	78.9	108	1	L5HDL
14	30	78.9	232	2	C84706
15	30	78.9	332	2	A45710
16	30	78.9	361	2	T30402
17	30	78.9	371	1	E64821
18	30	78.9	371	1	A85594
19	30	78.9	420	2	B86217
20	30	78.9	462	1	S35534
21	30	78.9	522	2	D82195
22	30	78.9	555	1	S24061
23	30	78.9	620	1	S3253
24	30	78.9	664	2	T01368
25	30	78.9	885	1	A55453
26	29	76.3	98	2	S36068
27	29	76.3	106	2	S38495
28	29	76.3	111	2	S47009
29	29	76.3	112	2	S51125

30	29	76.3	117	2	S04525	Ig lambda chain pr
31	29	76.3	132	2	A139004	Cyclin-dependent k
32	29	76.3	195	2	A27241	helic-destabilizin
33	29	76.3	202	1	MFVNAV	matrix protein - r
34	29	76.3	202	1	MFVNAV	matrix protein - r
35	29	76.3	202	1	MFVNAV	matrix protein - r
36	29	76.3	202	1	MFVNAV	matrix protein - r
37	29	76.3	232	2	S25742	Ig lambda chain -
38	29	76.3	243	2	A96744	hypothetical prote
39	29	76.3	312	1	S22397	pyruvate synthase
40	29	76.3	312	2	E84268	hypothetical prote
41	29	76.3	318	2	H69959	hypothetical prote
42	29	76.3	320	1	DDRT	helic-destabilizin
43	29	76.3	320	1	A44485	heterogeneous ribo
44	29	76.3	320	1	S02061	heterogeneous ribo
45	29	76.3	320	2	S30192	heterogeneous ribo

ALIGNMENTS

RESULT 1

L6HUA
Ig lambda chain V-VI region (AR) - human (tentative sequence)

C:Species: Homo sapiens (man)

C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 31-Mar-2000

C:Accession: A01987

R:Stetten, K.; Natvig, J.B.; Husby, G.; Juul, J.

Biochem. J. 195, 561-572, 1981

A:Title: The complete amino acid sequence of a prototype immunoglobulin-lambda light-

A:Reference number: A01987; MUID:82091000

A:Contents: amyloid protein AR

A:Accession: A01987

A:Molecule type: protein

A:Residues: 1112<SLK>

A:Note: About half of the lambda chain C region is missing from this protein

C:Comment: This protein was isolated from the spleen of a patient with amyloidosis.

C:Genetics:

A:Gene: GDB:IGLV6

A:Cross-references: GDB:119342; OMIM:147240

A:Map position: 22q11.2-22q11.2

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: amyloid; heterotetramer; immunoglobulin

F:15-93/Domain: immunoglobulin homology<IMK>

F:22-91/Disulfide bonds: #status predicted

Query Match 92.1% Score 35; DB 1; Length 112;

Best Local Similarity 85.7% Pred. No. 1.7;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNQRPS 7

DB 51 DDNQRPS 57

RESULT 2

L6HUA

Ig lambda chain V-I region (Wah) - human

C:Species: Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 02-Sep-1997

C:Accession: A01967

R:Takahashi, Y.; Takahashi, N.; Tetaert, D.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 80, 3686-3690, 1983

A:Title: Complete covalent structure of a human immunoglobulin D: sequence of the lam

A:Reference number: A01967; MUID:83221661

A:Accession: A01967

A:Molecule type: protein

A:Residues: 1-109<TAK>

C:Genetics:

A:Gene: GDB:IGLV6

A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 14 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-97/Region: V segment
F:15-91/Domain: Immunoglobulin homology <IMM>
F:98-109/Region: J segment
F:22-89/Disulfide bonds: #status predicted

Query Match 89.5%; Score 34; DB 1; Length 109;
Best Local Similarity 85.7%; Pred. No. 2.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNRPS 7
|||||
Db 51 EDNRPS 57

RESULT 3
LHUNM
Ig lambda chain V-I region (New) - human
C:Species: Homo sapiens (man)
C>Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 24-Nov-1999
C:Accession: A01964
R:Langer, B.; Steinmetz-Kayne, M.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 349, 945-951, 1968
A>Title: The complete amino acid sequence of Bence-Jones protein New (lambda type). Subg
A:Reference number: A01964; MUID:69060892
A:Accession: A01964
A:Molecule type: protein
A:Residues: 1-111 <LAN>
C:Comment: This is a Bence Jones protein.
C:Genetics:
A:Gene: GDB:IGLVE
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 14 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: blocked amino end; heterotetramer
F:15-91/Domain: Immunoglobulin homology <IMM>
F:1Modified site: blocked amino end (Gln) (probably pyrrolidone carboxylic acid) #statu
F:22-89/Disulfide bonds: #status predicted

Query Match 89.5%; Score 34; DB 1; Length 111;
Best Local Similarity 85.7%; Pred. No. 2.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNRPS 7
|||||
Db 51 EDNRPS 57

RESULT 4
L6HUB
Ig lambda chain precursor V-VI region (EB4) - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
C:Accession: A01990
R:Anderson, M.L.M.; Brown, L.; McKenzie, E.; Kellow, J.E.; Young, B.D.
Nucleic Acids Res. 13, 2931-2941, 1985
A>Title: Cloning and sequence analysis of an Ig lambda light chain mRNA expressed in the
A:Reference number: A01990; MUID:85215660
A:Accession: A01990
A:Molecule type: mRNA
A:Residues: 1-131 <AND>
C:Genetics:
A:Gene: GDB:IGLVE
A:Cross-references: GDB:119342; OMIM:147240

A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 14 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-131/Product: Ig lambda chain V-VI region (EB4) #status predicted <MAR>
F:20-41/Region: framework 1
F:34-112/Domain: Immunoglobulin homology <IMM>
F:42-54/Region: complementarity-determining 1
F:55-69/Region: framework 2
F:70-76/Region: complementarity-determining 2
F:77-110/Region: framework 3
F:111-118/Region: complementarity-determining 3
F:119-131/Region: framework 4
F:41-110/Disulfide bonds: #status predicted

Query Match 89.5%; Score 34; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNRP 6
|||||
Db 70 EDNRP 75

RESULT 5
S72947
hypothetical protein B235_F2_77 - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
C:Accession: S72947
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B235.
A:Reference number: S72587
A:Accession: S72947
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <SM>
A:Cross-references: EMBL:U00019; NID:q467079; PIDN:AA17283.1; PID:q467100

Query Match 89.5%; Score 34; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNRP 6
|||||
Db 103 EDNRP 108

RESULT 6
T02558
hypothetical protein AC2932590 (imported) - Arabidopsis thaliana
N:Alternate names: hypothetical protein T26B15.15
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 02-Feb-2001
C:Accession: T02558; AB4735
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.
A:Reference number: Z14678
A:Accession: T02558
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-704 <ROU>
A:Cross-references: EMBL:AC004681; NID:g3298532; PID:g3298547
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: A84420; MUID:20083487
 A:Accession: A84735
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-704 <STO>
 A:Cross-references: GB:AE002093; NID:g3298547; PIDN:AAC25941.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: ATG32590; T26B15.15
 A:Map position: 2
 A:Introns: 89/3; 123/3; 163/1; 184/3; 205/1; 400/3; 449/1; 543/1; 594/3; 631/3

Query Match 89.5%; Score 34; DB 2; Length 704;
 Best Local Similarity 85.7%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDNORPS 7
 Db 330 EDNORPS 336

RESULT 7
 LGHULT
 Ig lambda chain V-VI region (WLT) - human
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
 C:Accession: A01989
 R:Dwulet, F.E.; Strako, K.; Benson, M.D.
 Scand. J. Immunol. 22, 653-660, 1985
 A:Title: Amino acid sequence of a lambda VI primary (AI) amyloid protein (WLT).
 A:Reference number: A01989; MUID:66122667
 A:Accession: A01989
 A:Molecule type: protein
 A:Residues: 1-111 <DMW>
 C:Genetics:
 A:Gene: GDB:IGLV6
 A:Cross-references: GDB:119342; OMIM:147240
 A:Map position: 22q11.2-22q11.2
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:1-22/Region: framework 1
 F:15-93/Domain: immunoglobulin homology <IMW>
 F:23-35/Region: complementarity-determining 1
 F:36-50/Region: framework 2
 F:51-57/Region: complementarity-determining 2
 F:58-91/Region: framework 3
 F:92-101/Region: complementarity-determining 3
 F:102-111/Region: framework 4
 F:22-91/Disulfide bonds: #status predicted

Query Match 86.8%; Score 33; DB 1; Length 111;
 Best Local Similarity 85.7%; Pred. No. 4.5;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
 Db 51 EDNORPS 57

RESULT 8
 LGHUST
 Ig lambda chain V-VI region (SUR) - human
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 02-Sep-1997
 C:Accession: A01988
 R:Solomon, A.; Kyle, R.A.; Frangione, B.
 In Amyloidosis, Glenner, G.G.; Osseman, E.F.; Benditt, E.P.; Calkins, E.; Cohn, A.S., &
 A:Title: Light chain variable region subgroups of monoclonal immunoglobulins in amyloid

A:Reference number: A01988
 A:Accession: A01988
 A:Molecule type: Protein
 A:Residues: 1-111 <SOL>
 C:Genetics:
 A:Gene: GDB:IGLV6
 A:Cross-references: GDB:119342; OMIM:147240
 A:Map position: 22q11.2-22q11.2
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: Immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:1-22/Region: framework 1
 F:15-93/Domain: immunoglobulin homology <IMW>
 F:23-35/Region: complementarity-determining 1
 F:36-50/Region: framework 2
 F:51-57/Region: complementarity-determining 2
 F:58-91/Region: framework 3
 F:92-100/Region: complementarity-determining 3
 F:101-111/Region: framework 4
 F:22-91/Disulfide bonds: #status predicted

Query Match 84.2%; Score 32; DB 1; Length 111;
 Best Local Similarity 85.7%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EDNORPS 7
 Db 51 EDORPS 57

RESULT 9
 T27770
 hypothetical protein ZK177.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T27770
 R:Anderson, K.
 submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of C. elegans cosmid ZK177.
 A:Reference number: Z20416
 A:Accession: T27770
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-460 <AND>
 A:Cross-references: EMBL:U01321; PIDN:AAB36965.1; GSPDB:GN00020; CESP:ZK177.1
 A:Experimental source: strain Bristol N2; clone ZK177
 C:Genetics:
 A:Gene: CESP:ZK177.1
 A:Map position: 2
 A:Introns: 20/1; 31/3; 71/3; 111/1; 127/1; 294/1; 323/1
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein ZK177.1

Query Match 84.2%; Score 32; DB 2; Length 460;
 Best Local Similarity 71.4%; Pred. No. 32;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
 Db 153 QDNORPA 159

RESULT 10
 T20690
 hypothetical protein F10B5.3 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T20690
 R:Sim, M.
 submitted to the EMBL Data Library, February 1995
 A:Reference number: Z19310

A:Accession: T20690
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-471 <MIL>
 A:Cross-references: EMBL:Z48334; PIDN:CA88310.1; GSPDB:GN00020; CESP:FI085.3
 A:Experimental source: clone F1085
 C:Genetics:
 A:Gene: CESP:FI085.3
 A:Map position: 2
 A:Introns: 36/1; 114/3; 319/1; 362/2

Query Match 84.2%; Score 32; DB 2; Length 471;
 Best Local Similarity 85.7%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDNORPS 7
 |||||
 Db 3 EDNORPS 9

RESULT 11
 PC4283
 anti-SS-A/Ro 60k peptide light chain E-60 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 16-Jul-1999
 C:Accession: PC4283
 R:Suzuki, H.; Takemura, H.; Suzuki, M.; Sekine, Y.; Kashiwagi, H.
 Biochem. Biophys. Res. Commun. 232, 101-106, 1997
 A:Title: Molecular cloning of anti-ss-A/Ro 60-kDa peptide fab fragments from infiltratin
 A:Reference number: PC4279; MUID:97236289
 A:Accession: PC4283
 A:Molecule type: protein
 A:Residues: 1-107 <SUZ>
 C:Comment: This antibody is commonly found in systemic autoimmune diseases such as Sjogren
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 F:13-87/Domain: immunoglobulin homology <IMM>

Query Match 81.6%; Score 31; DB 2; Length 107;
 Best Local Similarity 71.4%; Pred. No. 12;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
 |||||
 Db 47 QDNKRPS 53

RESULT 12
 E81447
 transcription elongation factor Cj0287c [Imported] - Campylobacter jejuni (strain NCTC 1
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
 C:Accession: E81447
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barré
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h
 A:Reference number: A81250; MUID:20150912
 A:Accession: E81447
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-161 <PAR>
 A:Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CA872754.1; PID:g696776
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: greA; Cj0287c
 C:Superfamily: transcription elongation factor greB
 C:Keywords: transcription factor

Query Match 81.6%; Score 31; DB 2; Length 161;
 Best Local Similarity 71.4%; Pred. No. 18;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
 |||||
 Db 22 KDNORPA 28

RESULT 13
 L5HDDL
 Ig lambda chain V-V region (Del) - human (tentative sequence)
 C:Species: Homo sapiens (man)
 C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000
 C:Accession: A01985
 R:Eulitz, M.
 Eur. J. Biochem. 50, 49-69, 1974
 A:Title: A new subgroup of human L-chains of the lambda-type. Primary structure of Be
 A:Reference number: A01985; MUID:75112179
 A:Accession: A01985
 A:Molecule type: protein
 A:Residues: 1-108 <EUL>
 A:Note: this is the first sequenced V region of lambda chain subgroup V
 C:Comment: This is a Bence Jones protein.
 C:Genetics:
 A:Gene: GDB:IGLV6
 A:Cross-references: GDB:119342; OMIM:147240
 A:Map position: 22q11.2-22q11.2
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (h
 main disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer
 F:14-88/Domain: immunoglobulin homology <IMM>
 F:21-86/Disulfide bonds: #status predicted

Query Match 78.9%; Score 30; DB 1; Length 108;
 Best Local Similarity 71.4%; Pred. No. 20;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDNORPS 7
 |||||
 Db 48 EDNDRA 54

RESULT 14
 C84706
 probable small nuclear ribonucleoprotein U2B [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: C84706
 R:Llin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: C84706
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-232 <STO>
 A:Cross-references: GB:AE002093; NID:g2347192; PIDN:AC16931.1; GSPDB:GN00139
 C:Superfamily: small nuclear ribonucleoprotein U1A; ribonucleoprotein repeat homology

Query Match 78.9%; Score 30; DB 2; Length 232;
 Best Local Similarity 71.4%; Pred. No. 44;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
 |||||
 Db 119 EDSORPN 125

RESULT 15
A45710
BRRF2 gene homolog - bovine herpesvirus 4 (fragment)
C:Species: bovine herpesvirus 4
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: A45710
R:Van Santen, V. L.
J. Virol. 67, 773-784, 1993
A:Title: Characterization of a bovine herpesvirus 4 immediate-early RNA encoding a homolog
A:Reference number: A45710; NCID:93124571
A:Accession: A45710
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-332 <VAN>
A:Cross-references: GB:I01099; NID:930760; PIDN:AAA46056.1; PID:930763
A:Note: sequence extracted from NCBI backbone (NCBIN:122331, NCBIP:122333)

Query Match 78.9%; Score 30; DB 2; Length 332;
Best Local Similarity 83.3%; Pred. NO. 63;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNORP 6
1:||||
DB 11 EENORP 16

Search completed: November 19, 2001, 08:14:23
Job time: 196 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:35 ; Search time 27.32 seconds

(without alignments)
8.777 Million cell updates/sec

Title: US-09-610-118-65

Sequence: 38
1 EDNORPS 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	35	92.1	112 1	LV6A_HUMAN
2	34	89.5	109 1	LV1F_HUMAN
3	34	89.5	111 1	LV1C_HUMAN
4	34	89.5	131 1	LV6E_HUMAN
5	33	86.8	111 1	LV6D_HUMAN
6	32	84.2	111 1	LV6C_HUMAN
7	32	84.2	108 1	KS41_CAEEL
8	30	78.9	108 1	LV5A_HUMAN
9	30	78.9	332 1	YIE2_HSVB4
10	30	78.9	371 1	YL1I_ECOLI
11	30	78.9	551 1	ETV4_HUMAN
12	30	78.9	620 1	ETV4_MOUSE
13	30	78.9	195 1	ITK_HUMAN
14	29	76.3	195 1	ROA1_BOVIN
15	29	76.3	202 1	VMAT_RABVA
16	29	76.3	202 1	VMAT_RABVC
17	29	76.3	202 1	VMAT_RABVE
18	29	76.3	202 1	VMAT_RABVP
19	29	76.3	202 1	VMAT_RABVU
20	29	76.3	318 1	YOH0_BACSU
21	29	76.3	319 1	ROA1_MACOU
22	29	76.3	319 1	ROA1_MOUSE
23	29	76.3	319 1	ROA1_RAT
24	29	76.3	371 1	ROA1_HUMAN
25	29	76.3	475 1	PRTG_ERWCH
26	29	76.3	477 1	ETV1_HUMAN
27	29	76.3	477 1	ETV1_MOUSE
28	29	76.3	510 1	ERM_HUMAN
29	29	76.3	578 1	EDNORPS
30	29	76.3	871 1	SYA_AQUPY
31	29	76.3	924 1	ORC1_DROME
32	29	76.3	1131 1	PHYA_SORBI
33	29	76.3	1140 1	RA18_SCHPO

34	29	76.3	1193 1	DPOL_ADE04	P87503 human adeno
35	29	76.3	1239 1	V120_EBV	P03189 Epstein-Barr
36	29	76.3	1682 1	C1N6_HUMAN	P01118 human saplen
37	29	76.3	2164 1	POIG_HRV89	P07210 human rhino
38	28	73.7	78 1	VXIS_BPMD2	O22002 mycobacteri
39	28	73.7	107 1	LV4C_HUMAN	P01717 homo saplen
40	28	73.7	207 1	RCSA_ECOLI	P24210 Escherichia
41	28	73.7	446 1	NK0X_ENTFA	P37061 enterococu
42	28	73.7	449 1	CMGA_BOVIN	P05059 bos taurus
43	28	73.7	486 1	UAP1_CANAL	O74933 candida alb
44	28	73.7	562 1	AMAI_PLAER	P22622 plasmodium
45	28	73.7	817 1	VRP1_YEAST	P37370 Saccharomyc

ALIGNMENTS

RESULT 1
LV6A_HUMAN STANDARD: PRT: 112 AA.

AC P01721;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG LAMBDA CHAIN V-VI REGION AR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]

SEQUENCE (AMYL0ID PROTEIN AR).
RX MEDLINE=82091000; PubMed=6797401;
RA Sletten K., Navvig J.B., Husby G., Juul J.;
RT "The complete amino acid sequence of a prototype
PT immunoglobulin-lambda light-chain-fibrin protein AR.";
RL Biochem. J. 195;561-572(1981).
CC -1- MISCELLANEOUS: ABOUT HALF OF THE LAMBDA CHAIN C REGION IS MISSING
CC FROM THIS PROTEIN.
CC -1- MISCELLANEOUS: THIS PROTEIN WAS ISOLATED FROM THE SPLEEN OF A
CC PATIENT WITH AMYLOIDOSIS.
CC PIR: A01987; L6HUR.
DR HSSP: P01709; 2MCG.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region; Amyloid.
FT NON_TER 112 112
SQ SEQUENCE 112 AA: 11918 MW: 570BCD9A368E1FE CRC64:

Query Match 92.1%; Score 35; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 0.55;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
DB 51 EDNORPS 57

RESULT 2
LV1F_HUMAN STANDARD: PRT: 109 AA.
AC P04208;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG LAMBDA CHAIN V-I REGION WAH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=83221661; PubMed=6407018;

RA Takahashi Y., Takahashi N., Tetaert D., Putnam F.W.;
 RT "Complete covalent structure of a human immunoglobulin D: sequence of
 RT the lambda light chain."
 RT Proc. Natl. Acad. Sci. U.S.A. 80:3686-3690(1983).
 DR PIR: A01967; L1HUMA.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 97 V SEGMENT.
 FT DOMAIN 98 109 J SEGMENT.
 FT DISULFID 22 89 BY SIMILARITY.
 FT NON_TER 109 109
 SO SEQUENCE 109 AA: 11725 MW: 817785F68D9BAC CRC64:

Query Match
 Best Local Similarity 89.5%; Score 34; DB 1; Length 109;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
 DB 51 EDNORPS 57

RESULT 3
 LV6D_HUMAN STANDARD; PRT; 111 AA.
 ID LV6D_HUMAN
 AC P01701;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG LAMBDA CHAIN V-I REGION NEW.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=69060892; PubMed=4177823;
 RA Langer B., Steinmetz-Kayne M., Hilschmann N.;
 RT "The complete amino acid sequence of Bence Jones protein New (lambda-
 RT type). Subgroups in the variable part of immunoglobulin L-chains of
 RT the lambda-type."
 RL Hoppe-Seiler's Z. Physiol. Chem. 349:945-951(1968).
 CC -1- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.
 DR PIR: A01964; L1HUMA.
 DR HSSP: P01703; 7FAB.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region; Bence-Jones protein.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 22 89 BY SIMILARITY.
 FT NON_TER 111 111
 SO SEQUENCE 111 AA: 11453 MW: AAECBCA3C49F2AD3 CRC64:

Query Match
 Best Local Similarity 89.5%; Score 34; DB 1; Length 111;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
 DB 51 EDNORPS 57

RESULT 4
 LV6E_HUMAN STANDARD; PRT; 131 AA.
 ID LV6E_HUMAN
 AC P06319;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG LAMBDA CHAIN V-VI REGION EB4 PRECURSOR.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215660; PubMed=3923440;
 RA Anderson M.L.M., Brown L., McKenzie E., Kellow J.E., Young B.D.;
 RT "Cloning and sequence analysis of an Ig lambda light chain mRNA
 RT expressed in the Burkitt's lymphoma cell line EB4."
 RL Nucleic Acids Res. 13:2931-2941(1985).
 DR PIR: A01990; L6HUB.
 DR HSSP: P01709; 2MCG.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 131 IG LAMBDA CHAIN V-VI REGION EB4.
 FT DOMAIN 20 41 FRAMEWORK 1.
 FT DOMAIN 42 54 COMPLEMENTARITY-DETERMINING 1.
 FT DOMAIN 55 69 FRAMEWORK 2.
 FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING 2.
 FT DOMAIN 77 110 FRAMEWORK 3.
 FT DOMAIN 111 118 COMPLEMENTARITY-DETERMINING 3.
 FT DOMAIN 119 131 FRAMEWORK 4.
 FT DISULFID 41 110 BY SIMILARITY.
 FT NON_TER 131 131
 SO SEQUENCE 131 AA: 14147 MW: 02A9179C8C05C2CD CRC64:

Query Match
 Best Local Similarity 89.5%; Score 34; DB 1; Length 131;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORP 6
 DB 70 EDNORP 75

RESULT 5
 LV6D_HUMAN STANDARD; PRT; 111 AA.
 ID LV6D_HUMAN
 AC P06318;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG LAMBDA CHAIN V-VI REGION WLT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86122667; PubMed=4089539;
 RA Dwyer F.E., Strako K., Benson M.D.;
 RT "Amino acid sequence of a lambda VI primary (VL) amyloid protein
 RT (WLT)."
 RL Scand. J. Immunol. 22:653-660(1985).
 DR PIR: A01989; L6HULT.
 DR HSSP: P01709; 2MCG.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 22
 FT DOMAIN 23 35 COMPLEMENTARITY-DETERMINING 1.
 FT DOMAIN 36 50 FRAMEWORK 2.
 FT DOMAIN 51 57 COMPLEMENTARITY-DETERMINING 2.
 FT DOMAIN 58 91 FRAMEWORK 3.
 FT DOMAIN 92 101 COMPLEMENTARITY-DETERMINING 3.
 FT DOMAIN 102 111 FRAMEWORK 4.
 FT DISULFID 22 91 BY SIMILARITY.
 FT NON_TER 111 111
 SO SEQUENCE 111 AA: 11966 MW: 0C88B2FE37BCE24F CRC64:

Query Match 86.8%; Score 33; DB 1; Length 111;
Best Local Similarity 85.7%; Pred. No. 1.5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
DB 51 EDNORPS 57

RESULT 6
LV6C_HUMAN STANDARD; PRT; 111 AA.
AC P06317;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG LAMBDA CHAIN V-VI REGION SUT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RA Solomon A., Kyle R.A., Frangione B.;
RT *Light chain variable region subgroups of monoclonal immunoglobulins
in amyloidosis AL.*;
RL (In) Glenner G.G., Osseman E.F., Benditt E.P., Calkins E.,
RL Cohen A.S., Zucker-Franklin D. (eds.);
RL Amyloidosis, pp.449-462, Plenum Press, New York (1986).
DR PIR: A01988; L6HJST.
DR HSSP: P01703; 7FAB.
DR InterPro: IPR003006; -
KW Immunoglobulin V region.
FT DOMAIN 1 22 FRAMEWORK 1.
FT DOMAIN 2 35 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 3 50 FRAMEWORK 2.
FT DOMAIN 4 57 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 5 91 FRAMEWORK 3.
FT DOMAIN 6 92 COMPLEMENTARITY-DETERMINING 3.
FT DISULFID 101 111 FRAMEWORK 4.
FT DISULFID 22 91 BY SIMILARITY.
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12247 MW; 0941DD547D983598 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 111;
Best Local Similarity 85.7%; Pred. No. 2.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDNORPS 7
DB 51 EDNORPS 57

RESULT 7
YS41_CAEEL STANDARD; PRT; 460 AA.
AC Q09370;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOPHYSICAL 52.5 KDA PROTEIN ZK177.1 IN CHROMOSOME 11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RP SMRAIN-BRISTOL N2;

RA Anderson K.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U21321; MAG00044.1; -
DR WormPep: ZK177.1; CE02089.
KW Hypothetical protein.
SQ SEQUENCE 460 AA; 52498 MW; F5E295BCE6CC6F58 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 460;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
DB 153 QDNORPA 159

RESULT 8
LV5A_HUMAN STANDARD; PRT; 108 AA.
AC P01719;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG LAMBDA CHAIN V-V REGION DEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RA MEDLINE=75112179; PubMed=4452363;
RX Eulitz M.;
RT "A new subgroup of human I-chains of the lambda-type. Primary
RT structure of Bence-Jones protein DEL.*";
RL Eur. J. Biochem. 50:49-69(1974).
CC -1- MISCELLANEOUS: THIS IS THE FIRST SEQUENCED V REGION OF LAMBDA
CHAIN SUBGROUP V.
CC -1- MISCELLANEOUS: THIS IS A BENGE-JONES PROTEIN.
DR PIR: A01985; LSHUDL.
DR HSSP: P01703; 7FAB.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11342 MW; B8E8ED9C09C9E451 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 108;
Best Local Similarity 71.4%; Pred. No. 6.7;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDNORPS 7
DB 48 EDNDRA 54

RESULT 9
YIE2_HSVB4 STANDARD; PRT; 332 AA.
AC Q02484;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)

```

DE HYPOTHETICAL PROTEIN IN IE2 5' REGION (FRAGMENT).
OS Bovine herpesvirus type 4 (isolate DN-599).
OC Viruses: dsDNA viruses, no RNA stage: Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=10355;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=93124571; PubMed=8380465;
RA Van Santen V.L.;
RT "Characterization of a bovine herpesvirus 4 immediate-early RNA
  encoding a homolog of the Epstein-Barr virus R transactivator.";
RL J. Virol. 67:773-784(1993).
CC -1- SIMILARITY: TO EBV BRRE2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L01099; AAA6056.1; -.
DR PIR: A45710; A45710.
KW Hypothetical protein.
FT NON_TER
FT SEQUENCE 332 AA; 37951 MW; E0B6FAF287C5315 CRC64;

Query Match          78.9%; Score 30; DB 1; Length 332;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORP 6
Db 11 ENORP 16

RESULT 10
YLII_ECOLI STANDARD; PRT; 371 AA.
ID YLII_ECOLI
AC P7804;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 41.1 KDA PROTEIN IN MOEA-DACC INTERGENIC REGION
DE PRECURSOR.
GN YLII.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RP SEQUENCE FROM N.A.
RX STRAIN=K12; PubMed=9278503;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN (12)
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Molt H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horinouchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome

```

```

RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: STRONG, TO SYNECHOCYSTIS PCC 6803 SUR1608 AND SOME, TO
CC A.CALCOBACTERIUS GLUCOSE DEHYDROGENASE-B (GDH).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF000186; AAC73924.1; -.
DR EMBL: D90722; BAA35540.1; -.
DR Ecogene: EG31480; Y111.
KW Hypothetical protein; Signal.
FT SIGNAL 1 20
FT CHAIN 21 371
FT SEQUENCE 371 AA; 41054 MW; 99DB08FA302F50B9 CRC64;

Query Match          78.9%; Score 30; DB 1; Length 371;
Best Local Similarity 71.4%; Pred. No. 27;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORPS 7
Db 166 ENORPT 172

RESULT 11
EYV4_HUMAN STANDARD; PRT; 551 AA.
ID EYV4_HUMAN
AC P43268;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ADENOVIRUS E1A ENHANCER BINDING PROTEIN (E1A-F) (ETS TRANSLLOCATION
DE VARIANT 4) (FRAGMENT).
GN EYV4 OR E1AF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=95293380; PubMed=7774926;
RA Friedlan L.S., Ostermeyer E.A., Lynch E.D., Szabo C.I., Meza J.E.,
RA Anderson L.A., Dowd P., Lee M.K., Rowell S.E., Ellison J.,
RA Boyd J., King M.C.;
RT "22 genes from chromosome 17q21: cloning, sequencing, and
RT characterization of mutations in breast cancer families and tumors.";
RL Genomics 25:256-263(1995).
RN (12)
RP SEQUENCE OF 90-551 FROM N.A.
RX MEDLINE=93181246; PubMed=8441666;
RA Higashino F., Yoshida K., Fujinaga K., Kamio K., Fujinaga K.;
RT "Isolation of a cDNA encoding the adenovirus E1A enhancer binding
RT protein: a new human member of the ets oncogene family.";
RL Nucleic Acids Res. 21:547-553(1993).
CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO THE ENHANCER OF
CC THE ADENOVIRUS E1A GENE; THE CORE-BINDING SEQUENCE IS
CC 5'(AC)GCA(TAT)GT-3'.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: U18018; AA95991.1; -
DR EMBL: D12765; BAA02234.1; -
DR HSSP: P14921; 2STW.
DR TRANSFAC: T00685; -
DR MIM: 600711; -
DR InterPro: IPR000418; -
DR Pfam: PF00178; Ets; 1.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
KW DNA-binding: Nuclear protein; Activator; Transcription regulation.
FT DOMAIN 1 142 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 215 311 GLN-RICH.
FT DNA_BIND 408 488 ETS-DOMAIN.
FT CONFLICT 91 94 GNGS -> EMSD (IN REF. 2).
SQ SEQUENCE 551 AA; 60570 MW; 9D649C8A1FBDB8DB CRC64;

```

```

Query Match 78.9%; Score 30; DB 1; Length 551;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 DNORPS 7
DB 501 DNORPA 506

```

```

RESULT 12
ETV4_MOUSE STANDARD; PRT; 555 AA.
ID ETV4_MOUSE
AC P28322;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE POLYOMAVIRUS ENHANCER ACTIVATOR 3 (PEA3 PROTEIN) (ETS TRANSLLOCATION
DE VARIANT 4).
DE ETV4 OR PEA3 OR PEA-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92192459; PubMed=1547944;
RA Xin J.-H., Cowie A., Lachance P., Hassell J.A.;
RT "Molecular cloning and characterization of PEA3, a new member of the
RT Ets oncogene family that is differentially expressed in mouse
RT embryonic cells.";
RL Genes Dev. 6:481-496(1992).
CC -1- FUNCTION: BINDS TO THE PEA3 MOTIF (5'-AGGAAG-3'). MAY PLAY A
CC REGULATORY ROLE DURING EMBRYOGENESIS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: EPIDIDYMIS AND THE BRAIN.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X63190; CAA44872.1; -
DR PIR: S24061; S24061.
DR HSSP: P14921; 2STW.
DR TRANSFAC: T00684; -
DR MCD: MGI:99423; PEA3.
DR InterPro: IPR000418; -

```

```

DR Pfam: PF00178; Ets; 1.
DR PRINTS: PR00454; ETSDOMAIN.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
KW DNA-binding: Nuclear protein; Activator; Transcription regulation;
KW Phosphorylation.
FT DOMAIN 124 150 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 217 315 GLN-RICH.
FT DNA_BIND 412 492 ETS-DOMAIN.
SQ SEQUENCE 555 AA; 60846 MW; 278D76BA4A8A4D4B2 CRC64;

```

```

Query Match 78.9%; Score 30; DB 1; Length 555;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 DNORPS 7
DB 505 DNORPA 510

```

```

RESULT 13
ITK_HUMAN STANDARD; PRT; 620 AA.
ID ITK_HUMAN
AC Q08881;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TYROSINE-PROTEIN KINASE ITK/TSK (EC 2.7.1.112) (T-CELL-SPECIFIC
DE KINASE) (TYROSINE-PROTEIN KINASE LYK) (KINASE EMT).
DE ITK OR LYK OR EMT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93279365; PubMed=8504851;
RA Tanaka N., Asao H., Ohtani K., Nakamura M., Sugamura K.;
RT "A novel human tyrosine kinase gene inducible in T cells by
RT interleukin 2.";
RL FEBS Lett. 324:1-5(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93372354; PubMed=8364206;
RA Gibson S., Leung B., Squire J.A., Hill M., Arima N., Goss P.,
RA Hogg D., Mills G.B.;
RT "Identification, cloning, and characterization of a novel human
RT T-cell-specific tyrosine kinase located at the hematopoietin complex
RT on chromosome 5q.";
RL Blood 82:1561-1572(1993).
CC -1- FUNCTION: PLAYS A ROLE IN T CELL PROLIFERATION AND
CC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: LOCALIZES TO CELL SURFACE RECEPTORS IN THE
CC PLASMA MEMBRANE AFTER STIMULATION WITH RESPECTIVE RECEPTORS (TCR,
CC CD28, CD2) IN T-CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: T-CELL LINES AND NATURAL KILLER CELL LINES.
CC -1- INDUCTION: BY INTERLEUKIN 2.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE BTK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----

```

```

CC -----
DR EMBL: D13720; BAA02873.1; -
DR EMBL: L10717; AAA36748.1; -
DR EMBL: S65186; AAB28072.2; -
DR PIR: S33253; S33253.
DR HSSP: Q03526; 1AMJ.
DR MIM: 186973; -.
DR InterPro: IPR000719; -.
DR InterPro: IPR000980; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR001452; -.
DR InterPro: IPR001562; -.
DR InterPro: IPR001849; -.
DR Pfam: PF00779; BTK; 1.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00069; PKINASE; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00402; TECHDOMAIN.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR Transferrase: Tyrosine-protein kinase; Phosphorylation;
KM ATP-binding; SH2 domain; SH3 domain.
FT DOMAIN 4 111
FT DOMAIN 171 231
FT DOMAIN 239 338
FT DOMAIN 363 615
FT NE_BIND 369 377
FT BINDING 391 391
FT ACT_SITE 482 482
FT MOD_RES 512 512
FT CONFLICT 331 331
SQ SEQUENCE 620 AA; 71831 MW; DA6396BD309319D CRC64;

Query Match 78.9%; Score 30; DB 1; Length 620;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORP 6
DB 162 EDNRRP 167

RESULT 14
ID ROAL_BOVIN STANDARD; PRT; 195 AA.
AC P09867;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1 (HELIX-DESTABILIZING
DE PROTEIN) (SINGLE-STRAND BINDING PROTEIN) (HNRNP CORE PROTEIN A1)
DE (UNWINDING PROTEIN 1) (UPL1) (FRAGMENT).
GN HNRPA1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP TISSUE=Thymus;
RC MEDLINE=87193538; PubMed=3032834;
RA Merrill B.M., Lopresti M.B., Stone K.L., Williams K.R.;
RT "Amino acid sequence of UPL1, an hnrnp-derived single-stranded nucleic
RT acid binding protein from calf thymus.";
```

```

RL Int. J. Pept. Protein Res. 29:21-39(1987).
RN [2]
RP SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=85298222; PubMed=2994041;
RA Williams K.R., Stone K.L., Lopresti M.B., Merrill B.M., Planck S.R.;
RT "Amino acid sequence of the UPL1 calf thymus helix-destabilizing
RT protein and its homology to an analogous protein from mouse
RT myeloma."
RL Proc. Natl. Acad. Sci. U.S.A. 82:5666-5670(1985).
RN [3]
RP SEQUENCE OF 1-184.
RC TISSUE=Thymus;
RX MEDLINE=86085929; PubMed=3941105;
RA Merrill B.M., Lopresti M.B., Stone K.L., Williams K.R.;
RT "High pressure liquid chromatography purification of UPL1 and UPL2, two
RT related single-stranded nucleic acid-binding proteins from calf
RT thymus."
RL J. Biol. Chem. 261:878-883(1986).
CC -1- FUNCTION: INVOLVED IN THE PACKAGING OF PRE-MRNA INTO HNRNP
CC PARTICLES, TRANSPORT OF POLY-A MRNA FROM THE NUCLEUS TO THE
CC CYTOPLASM AND MAY MODULATE SPLICING SITE SELECTION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. SHUTTLES CONTINUOUSLY BETWEEN THE
CC NUCLEUS AND THE CYTOPLASM ALONG WITH MRNA. COMPONENT OF
CC RIBONUCLEOSOMES.
CC -1- PTM: UPL1 IS DERIVED FROM A1 BY PROTEOLYTIC DEGRADATION.
CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
CC GLY-RICH PROTEINS.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
DR PIR: A27241; A27241.
DR PIR: A25626; A25626.
DR HSSP: P09651; 1UPL.
DR InterPro: IPR000504; -.
DR Pfam: PF00076; RRM; 2.
DR PROSITE: PS50102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; 2.
DR Nuclear protein; RNA-binding; Repeat; Ribonucleoprotein;
KM Methylation.
FT DOMAIN 3 93
FT DOMAIN 94 184
FT DOMAIN 13 96
FT DOMAIN 104 183
FT MOD_RES 1 1
FT MOD_RES 193 193
FT NON_TER 195 195
SQ SEQUENCE 195 AA; 22132 MW; 448090F547D67ED1 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 195;
Best Local Similarity 83.3%; Pred. No. 22;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORP 6
DB 92 EDORP 97

RESULT 15
ID VMAT_RABVA STANDARD; PRT; 202 AA.
AC P15200;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE MATRIX PROTEIN (M2 PHOSPHOPROTEIN).
GN M2.
OS Rabies virus (strain AVO1).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11293;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89150235; PubMed=3147698;
```

RA Poch O., Tordo N., Keth-G.;
 RT "Sequence of the 386 3' nucleotides of the genome of the AVO1 strain
 RT rabies virus: structural similarities in the protein regions involved
 RT in transcription.";
 RL Biochimie 70:1019-1029(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X13357; CAA31736.1; -
 DR PIR; S07816; MFVNAV.
 KW Phosphorylation; Matrix protein; Envelope protein.
 SQ SEQUENCE 202 AA; 23231 MW; DD139AE8BCFB2D2 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 202;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDNORPS 7
 |||
 Db 14 EDTOKPS 20

Search completed: November 19, 2001, 08:25:36
 Job time: 724 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:03 ; Search time 81.19 Seconds

(without alignments)
11.407 Million cell updates/sec

Title: US-09-610-118-65

Perfect score: 38

Sequence: 1 EDNORPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP unclassified:*
13: SP vertebrate:*
14: SP virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	89.5	266	2 049849	049849 mycobacteri
2	34	89.5	704	10 080896	080896 arabidopsis
3	33	86.8	478	5 043955	043955 leishmania
4	33	86.8	401	13 09PU01	09PU01 brachydanio
5	33	86.8	494	13 057586	057586 brachydanio
6	32	84.2	471	5 009534	009534 caenorhabdi
7	31	81.6	161	2 09PIK9	09PIK9 campylobact
8	31	81.6	738	5 09VAC9	09VAC9 drosophila
9	30	78.9	85	2 091987	091987 chlamydia t
10	30	78.9	232	10 023922	023922 arabidopsis
11	30	78.9	273	2 09R7R6	09R7R6 escherichia
12	30	78.9	336	13 09YHW5	09YHW5 gallus gall
13	30	78.9	355	2 09F3R6	09F3R6 streptomyce
14	30	78.9	361	14 09YMR9	09YMR9 lymphantria d
15	30	78.9	366	2 09RCR0	09RCR0 neisseria m
16	30	78.9	390	9 09FZR6	09FZR6 mycoplasma
17	30	78.9	409	10 09SGC6	09SGC6 arabidopsis
18	30	78.9	420	10 09SUF6	09SUF6 arabidopsis
19	30	78.9	522	2 085210	085210 vibrio chol

20	30	78.9	522	2 09K394	09K394 vibrio chol
21	30	78.9	664	10 064589	064589 arabidopsis
22	30	78.9	862	11 09R1R6	09R1R6 mus musculu
23	30	78.9	885	11 064610	064610 rattus norv
24	30	78.9	2473	11 09Q284	09Q284 mus musculu
25	29	76.3	76	14 088283	088283 snakehead r
26	29	76.3	94	14 088287	088287 snakehead r
27	29	76.3	112	14 082090	082090 human rhino
28	29	76.3	132	4 013195	013195 homo sapien
29	29	76.3	140	2 066306	066306 unidentified
30	29	76.3	173	4 013399	013399 homo sapien
31	29	76.3	179	4 016360	016360 homo sapien
32	29	76.3	202	14 09IPJ7	09IPJ7 rabies viru
33	29	76.3	236	13 09YHW7	09YHW7 gallus gall
34	29	76.3	311	1 059456	059456 halobacteri
35	29	76.3	312	1 09HOK0	09HOK0 halobacteri
36	29	76.3	317	4 095024	095024 homo sapien
37	29	76.3	328	10 092VX7	092VX7 arabidopsis
38	29	76.3	376	6 09TSV3	09TSV3 sus scrofa
39	29	76.3	400	10 09SOF7	09SOF7 brassica ju
40	29	76.3	463	2 09X143	09X143 thermotoga
41	29	76.3	473	11 092ZM0	092ZM0 mus musculu
42	29	76.3	477	13 093322	093322 xenopus lae
43	29	76.3	477	13 09YHW6	09YHW6 gallus gall
44	29	76.3	477	13 09W696	09W696 xenopus lae
45	29	76.3	510	11 09JHB1	09JHB1 mus musculu

ALIGNMENTS

RESULT 1
ID 049849 PRELIMINARY; PRT; 266 AA.
AC 049849;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)
DE B2235.F2_77.
OS Mycobacterium leprae.
OC Bacteria: Firmicutes; Actinobacteria: Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U00019; AAA17283.1; -;
SQ SEQUENCE 266 AA; 29602 MW; 02C50A1B4E36EC82 CRC64;

Query Match 89.5%; Score 34; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDNORP 6
Db 103 EDNORP 108
RESULT 2
ID 080896 PRELIMINARY; PRT; 704 AA.
AC 080896;
DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)
 DE PUTATIVE CONDENSIN PROTEIN.
 GN T26B15.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
 RA Brandon R.C., Sykes S.M., Mason T.M., Kellavage A.R., Adams M.D.,
 RA Somerville C.R., Venter J.C.;
 RT "Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC004681; AAC25941.1; -
 SO SEQUENCE 704 AA; 79017 MW; 69C8BFD56C040FA CRC64;

Query Match 89.5%; Score 34; DB 10; Length 704;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DNORPS 7
 DB 330 EENORPS 336

RESULT 3
 ID 043955 PRELIMINARY; PRT; 478 AA.
 AC 043955;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)
 DE NUPM1.
 GN TOR.
 OS Leishmania mexicana.
 OC Eukaryota; Eukleozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_Taxid=5665;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=LV 78;
 RX MEDLINE=98135664; PubMed=9476797;
 RA Deke S.;
 RT "Identification of a transcription factor like protein at the TOR
 RT locus in Leishmania mexicana amazonensis.";
 RL Mol. Biochem. Parasitol. 90:505-511(1997).
 DR EMBL: AF016581; AAC38851.1; -
 SO SEQUENCE 478 AA; 52237 MW; 60363DAE0210CEFA1 CRC64;

Query Match 86.8%; Score 33; DB 5; Length 478;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
 DB 342 DNORPS 347

RESULT 4
 ID 09PU01 PRELIMINARY; PRT; 491 AA.
 AC 09PU01;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE PEA3.
 GN Brachydanio rerio (Zebrafish) (Zebra danio).
 OS

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbortinae; Danio.
 OX NCBI_Taxid=7955;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC MEDLINE=20005800; PubMed=10534622;
 RA Munchberg S.R., Ober E.A., Steinbeisser H.;
 RT "Expression of the Ets transcription factors erm and pea3 in early
 RT zebrafish development.";
 RL Mech. Dev. 88:233-236(1999).
 DR EMBL: AF168008; AAD50434.1; -
 DR HSSP: P14921; 2STT.
 DR InterPro: IPR000418; -
 DR InterPro: IPR002341; -
 DR Pfam: PF00178; Ets; 1.
 DR PRINTS: PR00454; ETS_DOMAIN.
 DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
 DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
 DR PROSITE: PS0061; ETS_DOMAIN_3; 1.
 DR SMART: SM00413; ETS; 1.
 SO SEQUENCE 491 AA; 53476 MW; 11AB8A1813B3AE7B CRC64;

Query Match 86.8%; Score 33; DB 13; Length 491;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
 DB 440 DNORPS 445

RESULT 5
 ID 057586 PRELIMINARY; PRT; 494 AA.
 AC 057586;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE ETS-DOMAIN TRANSCRIPTION FACTOR PEA3.
 GN PEA3.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbortinae; Danio.
 OX NCBI_Taxid=7955;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC Brown L.A., Amores A., Schilling T.F., Jowett T., Baert J.L.,
 RA de Launolt Y., Sharrocks A.D.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ003200; CAA05980.1; -
 DR HSSP: P14921; 2STT.
 DR ZFIN: ZDB-GENE-990415-71; pea3.
 DR InterPro: IPR000418; -
 DR InterPro: IPR002341; -
 DR Pfam: PF00178; Ets; 1.
 DR PRINTS: PR00454; ETS_DOMAIN.
 DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
 DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
 DR PROSITE: PS0061; ETS_DOMAIN_3; 1.
 DR SMART: SM00413; ETS; 1.
 SO SEQUENCE 494 AA; 55620 MW; 81492553EA25C362 CRC64;

Query Match 86.8%; Score 33; DB 13; Length 494;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DNORPS 7
 DB 443 DNORPS 448

RESULT 6
ID 009534 PRELIMINARY; PRT; 471 AA.
AC 009534;
DT 01-JAN-1999 (TREMblrel. 09, Created)
DT 01-JAN-1999 (TREMblrel. 09, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
GN PUTATIVE 51.7 KDA ZINC FINGER PROTEIN F10B5.3 IN CHROMOSOME II.
OS F10B5.3.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL NZ;
RA Sims M.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
DR EMBL; Z48334; CAAB8310.1;
DR WormPep; F10B5.3; CE01545.
DR InterPro; IPR000210;
DR Pfam; PF00096; ZF-C2H2; 3.
DR PROSITE; PS50097; BTB; 1.
DR PROSITE; PS50028; ZINC_FINGER_C2H2; 1.
DR SMART; SM00355; ZNF_C2H2; 1.
KW Hypothetical protein; Zinc-finger; Metal-binding; Nuclear protein;
KW DNA-binding.
FT DOMAIN 435 440 POLY-SER.
FT ZN.FING 255 278 C2H2-TYPE.
SQ SEQUENCE 471 AA; 51751 MW; C9F7B0E3AEAB82CA CRC64;

Query Match
Best Local Similarity 84.2%; Score 32; DB 5; Length 471;
Matches 6; Conservative 85.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EDNRPS 7
Db 3 EDNRPS 9

RESULT 7
ID 09PIK9 PRELIMINARY; PRT; 161 AA.
AC 09PIK9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
GN TRANSCRIPTION ELONGATION FACTOR.
OS GRAA OR C10287.
OC Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_Taxid=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCTC 11168;
RX MEDLINE-20150912; Pubmed-10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Randleman M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.,
RT The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.
RL Nature 403:665-668(2000).
DR EMBL; AL139074; CAB72754.1;
DR InterPro; IPR001437;
DR Pfam; PF01272; GreA_Greb; 1.

DR PROSITE; PS00829; GREAB-1; 1.
DR PROSITE; PS00830; GREAB-2; UNKNOWN.1.
SQ SEQUENCE 161 AA; 18010 MW; 132C7DEC962A3239 CRC64;

Query Match
Best Local Similarity 81.6%; Score 31; DB 2; Length 161;
Matches 5; Conservative 71.4%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDNRPS 7
Db 22 KDNRPA 28

RESULT 8
ID 09VAC9 PRELIMINARY; PRT; 738 AA.
AC 09VAC9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
GN CG2052. PROTEIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE-20196006; Pubmed-10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Plamkoc H., Baldwin D.,
RA Baller R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Donnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Jellum B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mitsuhashi N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Reinartz K., Remington K., Saunders R.D.C., Saito V., Reese M.G.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Weissman D.A., Weissstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.N., Rubin G.M., Venter J.C.;
RT The genome sequence of Drosophila melanogaster.
RL Science 287:2185-2195(2000).
DR EMBL; AE003844; AAF59348.1;
DR HSSP; P08151; ZGLT.
DR Flybase; Fbgn0039905; CG2052.

DR InterPro: IPR000822; -.
 DR Pfam: PF00096; zfc-C2H2; 7.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 6.
 DR SMART: SM00355; ZNF_C2H2; 1.
 KW DNA-binding; Metal-binding; Zinc-finger.
 SO SEQUENCE 738 AA; 79951 MW; 5FE6C3E3AC44795C CRC64;

Query Match 81.6%; Score 31; DB 5; Length 738;
 Best Local Similarity 71.4%; Pred. No. 68;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDNORPS 7
 Db 730 QENORPS 736

RESULT 9
 Q919B7 PRELIMINARY; PRT; 85 AA.
 AC Q919B7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA588;
 RX MEDLINE=20085123; PubMed=10618077;
 RA Sturm-Ramirez K., Brumblay H., Dlop K., Gueye-Ndiaye A., Sankale J.L.,
 RA Thior I., N'Doye I., Hsieh C.C., Mboup S., Kanki P.J.;
 RT "Molecular epidemiology of genital Chlamydia trachomatis infection in
 high-risk women in Senegal, West Africa";
 RL J.Clin.Microbiol.38:138-145(2000).
 DR EMBL: AF178260; AAF26695.1; JOINED.
 DR EMBL: AF178257; AAF26695.1; JOINED.
 DR EMBL: AF178258; AAF26695.1; JOINED.
 DR EMBL: AF178259; AAF26695.1; JOINED.
 DR InterPro: IPR000604; -.
 DR Prodom: PD001717; -. 1.
 FT NON_TER 1 1
 FT NON_TER 85 85
 SQ SEQUENCE 85 AA; 8551 MW; A183857BB2CB076B CRC64;

Query Match 78.9%; Score 30; DB 2; Length 85;
 Best Local Similarity 71.4%; Pred. No. 14;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EDNORPS 7
 Db 20 EDNENPS 26

RESULT 10
 022922 PRELIMINARY; PRT; 232 AA.
 ID 022922;
 AC 022922;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE SPICEOSOMAL PROTEIN U2B ISOLOG.
 GN T09D09.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
 RA Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R.,
 RA Venter J.C.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC002338; AAC16931.1; -.
 DR HSSP: P08579; 1A9N.
 DR Mendel: 26539; Arch; 1206; 26539.
 DR InterPro: IPR000504; -.
 DR InterPro: IPR000834; -.
 DR Pfam: PF00076; rim; 2.
 DR PROSITE: PS00133; CARBOXYPEPT_2N_2; UNKNOWN_1.
 DR SMART: SM00360; RRM; 1
 SO SEQUENCE 232 AA; 26237 MW; A0066B07B3084E99 CRC64;

Query Match 78.9%; Score 30; DB 10; Length 232;
 Best Local Similarity 71.4%; Pred. No. 37;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDNORPS 7
 Db 119 EDSQRPN 125

RESULT 11
 Q9R7R6 PRELIMINARY; PRT; 273 AA.
 ID Q9R7R6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE ORF_ID=O209#7 (FRAGMENT).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajinara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horluchi T.;
 RT "A 718-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding
 to the 12.7-28.0 min Region on the Linkage Map";
 RL DNA Res. 3:137-155(1996).
 DR EMBL: D90721; BAA35532.1; -.
 FT NON_TER 273 273
 SQ SEQUENCE 273 AA; 30059 MW; E27C16D648123C6D CRC64;

Query Match 78.9%; Score 30; DB 2; Length 273;
 Best Local Similarity 71.4%; Pred. No. 43;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDNORPS 7
 Db 166 ENNORPT 172
 RESULT 12
 09YHW5 PRELIMINARY; PRT; 336 AA.
 ID 09YHW5;
 AC 09YHW5;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE ETS DOMAIN PROTEIN (FRAGMENT).

GN PEAK
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9029846; PubMed=9814709;
 RA Lin J.H., Salto T., Anderson D.J., Lance-Jones C., Jessell T.M.,
 RA Arber S.;
 RT "Functionally related motor neuron pool and muscle sensory afferent
 RT subtypes defined by coordinate ETS gene expression.";
 RL Cell 95:393-407(1998).
 DR EMBL: AF075708; AAC97203.1; -.
 DR HSSP: P28324; IBC8.
 DR InterPro: IPR000418; -.
 DR Pfam: PF00178; Ets: 1.
 DR PRINTS: PR00454; ETSDOMAIN.
 DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
 DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
 DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
 DR SMART: SM00413; ETS; 1.
 FT NON_TER 1
 SQ SEQUENCE 336 AA; 38518 MW; 0507A0A11995E4BB CRC64;

Query Match 78.9%; Score 30; DB 13; Length 336;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 DNORPS 7
 Db 287 DNORPS 292

RESULT 13
 O9F3F6 PRELIMINARY; PRT; 355 AA.
 AC O9F3F6;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE PUTATIVE SECRETED PROTEIN.
 GN 2SCD4.18
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kiese H.M., Denepaite D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL392175; CAC08363.1; -.
 SQ SEQUENCE 355 AA; 37716 MW; 3DE06B00F46288BE CRC64;

Query Match 78.9%; Score 30; DB 2; Length 355;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDNRPS 7
 Db 31 EDNRPS 37

RESULT 14
 O9YMR9 PRELIMINARY; PRT; 361 AA.
 AC O9YMR9;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE LDORF-55 PEPTIDE.
 OS Lymantria dispar multicaud nuclear polyhedrosis virus (LdMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99124785; PubMed=9887315;
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J.M., Rohmann G.F.;
 RT "Sequence and analysis of the genome of a baculovirus pathogenic for
 RT Lymantria dispar.";
 RL Virology 253:17-34(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J., Rohmann G.F.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kuzio J.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF081810; AAC70240.1; -.
 SQ SEQUENCE 361 AA; 39665 MW; 76D2D1627EB9F402 CRC64;

Query Match 78.9%; Score 30; DB 14; Length 361;
 Best Local Similarity 71.4%; Pred. No. 57;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EDNRPS 7
 Db 318 QDNDPS 324

RESULT 15
 O9RGRO PRELIMINARY; PRT; 366 AA.
 AC O9RGRO;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE LCBA.
 GN LCBA.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Brewer N.E., Coulthart M.B., Tyler S.D.;
 RT "Analysis of the region encoding the putative capsular biosynthetic
 RT genes of the Neisseria meningitidis serogroup I.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF112478; AAF21950.1; -.
 SQ SEQUENCE 366 AA; 43281 MW; F7A95F3843D1B67E CRC64;

Query Match 78.9%; Score 30; DB 2; Length 366;
Best Local Similarity 83.3%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNQRP 6
DB 330 EDNQRP 335

Search completed: November 19, 2001, 08:25:04
Job time: 742 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:14:23 ; Search time 48.99 Seconds
(without alignments)
12.439 Million cell updates/sec

Title: US-09-610-118-66

Perfect score: 39

Sequence: 1 SYDSSNVV 8

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	680	2 T08080	probable myrosinase
2	33	84.6	730	2 F96559	hypothetical prote
3	33	84.6	934	2 C81298	probable formate d
4	33	84.6	1037	2 A60163	glycoprotein rib -
5	33	84.6	1945	2 T13937	plexin A - fruit f
6	32	82.1	112	1 L6H048	ig lambda chain V -
7	32	82.1	469	2 B37837	probable alpha-amy
8	32	82.1	571	2 S58356	pept protein - Sta
9	31	79.5	130	2 T20762	hypothetical prote
10	31	79.5	245	2 A30154	IGF receptor alpha
11	31	79.5	249	2 G84781	hypothetical prote
12	31	79.5	261	2 T15520	hypothetical prote
13	31	79.5	271	2 JC4584	insulin-like growt
14	31	79.5	374	2 T21513	hypothetical prote
15	31	79.5	413	2 A70795	hypothetical prote
16	31	79.5	450	2 T21515	hypothetical prote
17	31	79.5	783	2 B84823	probable isocamylas
18	31	79.5	876	2 B96693	probable receptor
19	31	79.5	963	2 S45167	chitin synthase (E
20	30	76.9	229	2 S49494	replication proteol
21	30	76.9	293	2 T23806	hypothetical prote
22	30	76.9	311	2 T17613	hypothetical prote
23	30	76.9	358	2 T15069	hypothetical prote
24	30	76.9	429	2 S63360	hypothetical prote
25	30	76.9	458	2 T16041	hypothetical prote
26	30	76.9	462	2 B96560	hypothetical prote
27	30	76.9	462	2 T52115	myrosinase-binding
28	30	76.9	624	1 VCNV87	p87 capsid protein
29	30	76.9	624	2 T10374	p87 capsid protein

30	30	76.9	625	2 A84787	hypothetical prote
31	30	76.9	630	2 J01894	hypothetical 74.1k
32	30	76.9	646	2 T02398	hypothetical prote
33	30	76.9	754	2 S60464	mes-3 protein - Ca
34	30	76.9	769	1 JUHULM	leukocyte adhesion
35	30	76.9	798	2 A40526	integrin beta-7 ch
36	30	76.9	803	2 B84931	DNA topoisomerase
37	30	76.9	806	2 A46271	integrin beta-7 ch
38	30	76.9	817	2 T51787	hypothetical prote
39	30	76.9	1358	2 T22695	hypothetical prote
40	29	74.4	197	2 F96799	hypothetical prote
41	29	74.4	214	2 D75576	similar to 'MADS b
42	29	74.4	295	2 H65118	oxidoreductase, ir
43	29	74.4	296	2 B85991	hypothetical adeni
44	29	74.4	317	2 S74030	probable methyltra
45	29	74.4	354	2 B83099	hypothetical prote

ALIGNMENTS

RESULT 1
T08080
probable myrosinase-binding protein - rape
N:Alternate names: jasmonate inducible protein
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T08080
R:Geshi, N.; Brandt, A.
Planta 204, 295-304, 1998
A:Title: Two jasmonate-inducible myrosinase-binding proteins from Brassica napus L. s
A:Reference number: Z16340; MUID:98192006
A:Accession: T08080
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: mRNA
A:Residues: 1-680 <GES>
A:Cross-references: EMBL:Y11483; NID:e1023101; PIDN:CAA72271.1; PID:e304150
A:Experimental source: cv. Global; isolate a4; young seedlings
A>Note: jasmonate inducible

Query Match 84.6% Score 33; DB 2; Length 680;
Best Local Similarity 71.4% Pred. No. 38;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 YDSSNVV 8
|||||:
Db 308 YDSSNII 314

RESULT 2
F96559
hypothetical protein F5F19.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96559
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marzita Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tello, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: F96559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-730 <STO>
A:Cross-references: GB:AE005173; NID:g4220464; PIDN:AAD12691.1; GSPDB:GNO0141
C:Genetics:

A:Gene: F5P19.6
A:Map position: 1

Query Match 84.6%; Score 33; DB 2; Length 730;
Best Local Similarity 71.4%; Pred. No. 41;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 SYDSNNV 8
Db 511 YDSSNII 517

RESULT 3
C81298
probable formate dehydrogenase (EC 1.2.1.2) large chain Cj1511c [similarity] - Campylobacter
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C:Accession: C81298
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912
A:Accession: C81298
A:Molecule type: DNA
A:Residues: 1-934 <PAR>
A:Cross-references: GB:ALJ39078; GB:ALJ11168; NID:96968723; PIDN:CAB73932.1; PID:9696893
A:Experimental source: serotype O2, strain NCYC 11168
A:Note: A translation exception was noted by the authors
A:Note: In Genbank entry ALJ39078, release 117.0, PIDN:CAB73932.1, the selenocysteine U
C:Genetics:
A:Gene: fdha: Cj1511c
C:Superfamily: formate dehydrogenase
C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein; molybdenum; molybdopterin; NAD;
F:57,60,64,92/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
F:94,182/Active site: Lys, His #status predicted
F:181/Binding site: molybdopterin guanine dinucleotide (Cys) (covalent) #status predicted
F:181/Modified site: selenocysteine #status predicted

Query Match 84.6%; Score 33; DB 2; Length 934;
Best Local Similarity 62.5%; Pred. No. 53;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSNNV 8
Db 714 YDSSNII 721

RESULT 4
A60163
glycoprotein IIb - rat
C:Species: Rattus sp. (rat)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 13-Sep-1998
C:Accession: A60163; B60163
R:Poncz, M.; Newman, P.J.
Blood 75, 1282-1289, 1990
A:Title: Analysis of rodent platelet glycoprotein IIb: evidence for evolutionarily conse
A:Reference number: A60163; MUID:90181615
A:Accession: A60163
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1037 <PON>
A:Accession: B60163
A:Molecule type: protein
A:Residues: 32-50 <PON2>
C:Comment: This protein is proteolytically processed into a heavy chain and a light chain
C:Superfamily: integrin alpha-2b chain
C:Keywords: calcium; disulfide bond; glycoprotein

Query Match 84.6%; Score 33; DB 2; Length 1037;

Best Local Similarity 87.5%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSNNV 8
Db 252 SYDSNNV 259

RESULT 5
T13937
plexin A - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C:Accession: T13937
R:Winderberg, M.L.; Noordermeer, J.N.; Tamagnone, L.; Comoglio, P.M.; Spriggs, M.K.; Tes
Cell 95, 903-916, 1998
A:Title: Plexin A is a neuronal semaphorin receptor that controls axon guidance.
A:Reference number: 217621; MUID:99091049
A:Accession: T13937
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1945 <MIN>
A:Cross-references: EMBL:AF106932; NID:94056673; PID:94056674; PIDN:AAD09425.1
C:Genetics:
A:Gene: plexA
A:Cross-references: FlyBase:FBgn0025741
A:Map position: 4
C:Function:
A:Description: may function as repellents during axon guidance
C:Keywords: cell adhesion; nerve

Query Match 84.6%; Score 33; DB 2; Length 1945;
Best Local Similarity 62.5%; Pred. No. 1-2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSNNV 8
Db 1266 SYESSNII 1273

RESULT 6
L6H048
Ig lambda chain V-VI region (Nlg-48) - human
C:Species: Homo sapiens (man)
C:Date: 28-Feb-1980 #sequence_revision 28-Feb-1980 #text_change 02-Sep-1997
C:Accession: A01991
R:Takanashi, N.; Takayasu, T.; Isobe, T.; Shinoda, T.; Okuyama, T.; Shimizu, A.
J. Biochem. 86, 1523-1535, 1979
A:Title: Comparative study on the structure of the light chains of human immunoglobul
A:Reference number: A01991; MUID:80094390
A:Accession: A01991
A:Molecule type: protein
A:Residues: 1-112 <TAK>
A:Note: This is the first sequenced V region of lambda chain subgroup VI
C:Comment: This is a Bence Jones protein.
C:Genetics:
A:Gene: GDB:IGLV6
A:Cross-references: GDB:119342; OMIM:147240
A:Map position: 22q11.2-22q11.2
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (h
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer
F:15-93/Domain: immunoglobulin homology <IMM>
F:22-91/Disulfide bonds: #status predicted

Query Match 82.1%; Score 32; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 8.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSNNV 7

DB 93 SYDSNMV 99

RESULT 7

B37837

Probable alpha-amylase (EC 3.2.1.1) - Clostridium acetobutylicum (fragment)

C:Species: Clostridium acetobutylicum

C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 22-Jun-1999

C:Accession: B37837

R:Gertscher, U.; Duerre, P.

J. Bacteriol. 172, 6907-6918, 1990

A:Title: Cloning, sequencing, and molecular analysis of the acetate decarboxylase gene

A:Reference number: A37837; MUID:91072241

A:Accession: B37837

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-469 <GER>

A:Cross-references: GB:M55392; GB:M34078; NID:q144708; PIDN:AAA63759.1; PID:q144709

C:Superfamily: alpha-amylase, subtilis type; alpha-amylase core homology

C:Keywords: glycosidase; hydrolase

Query Match 82.1%; Score 32; DB 2; Length 469;

Best Local Similarity 85.7%; Pred. No. 42;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSSNMV 8

DB 286 YDSSNMV 292

RESULT 8

S58356

Pept protein - Staphylococcus epidermidis

C:Species: Staphylococcus epidermidis

C:Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 02-Feb-2001

C:Accession: S58356

R:Meyer, C.; Blerbaum, G.; Heidrich, C.; Reis, M.; Suelting, J.; Iglesias-Wind, M.I.; Ken

Eur. J. Biochem. 232, 478-489, 1995

A:Title: Nucleotide sequence of the jantibiotic Pep5 biosynthetic gene cluster and funct

A:Reference number: S58356; MUID:96035882

A:Accession: S58356

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-571 <ME2>

A:Cross-references: EMBL:249865; NID:9945014; PIDN:CAA90021.1; PID:9945015

C:Genetics:

A:Gene: pepT

C:Superfamily: Escherichia coli ABC transporter mlaA; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop

F:348-543/Domain: ATP-binding cassette homology <ABC>

F:365-372/Region: nucleotide-binding motif A (P-loop)

Query Match 82.1%; Score 32; DB 2; Length 571;

Best Local Similarity 75.0%; Pred. No. 51;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YDSSNMV 8

DB 341 YDSSNMV 348

RESULT 9

T20762

hypothetical protein F11C1.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T20762

R:Palmer, S.

submitted to the EMBL Data Library, September 1995

A:Reference number: Z19321

A:Accession: T20762

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-130 <WIL>

A:Cross-references: EMBL:Z54270; PIDN:CAA91026.1; GSPDB:GN00028; CESP:F11C1.2

A:Experimental source: clone F11C1

C:Genetics:

A:Gene: CESP:F11C1.2

A:Map position: X

A:introns: 23/3; 63/3; 89/1

Query Match 79.5%; Score 31; DB 2; Length 130;

Best Local Similarity 85.7%; Pred. No. 17;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YDSSNMV 7

DB 68 SFDSSNV 74

RESULT 10

A30154

IgE receptor alpha chain precursor - rat

N:Alternate names: Fc-epsilon-R alpha chain precursor

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 21-Jan-2000

C:Accession: C31327; A31327; A30154; A27116; I55304

R:Jill, P.T.; Albrandt, K.; Robertson, M.W.

Proc. Natl. Acad. Sci. U.S.A. 85, 5639-5643, 1988

A:Title: cDNA heterogeneity suggests structural variants related to the high-affinity

A:Reference number: A94203; MUID:88289772

A:Accession: C31327

A:Molecule type: mRNA

A:Residues: 1-245 <LTI>

A:Cross-references: GB:M21622; GB:J03811

A:Experimental source: basophilic leukemia cell line, clone R3-4

A:Accession: A31327

A:Molecule type: mRNA

A:Residues: 21-245 <LTI>

A:Cross-references: GB:M21622; NID:q204109; PIDN:AAA41146.1; PID:q204110; GB:J03811

A:Experimental source: basophilic leukemia cell line

R:Shimizu, A.; Tepler, I.; Benfey, P.N.; Berenstein, E.H.; Siraganian, R.P.; Leder, P

Proc. Natl. Acad. Sci. U.S.A. 85, 1907-1911, 1988

A:Title: Human and rat mast cell high-affinity immunoglobulin E receptors: characteri

A:Reference number: A94191; MUID:88158102

A:Accession: A30154

A:Molecule type: mRNA

A:Residues: 1-245 <SHI>

A:Cross-references: GB:J03606; NID:q205331; PIDN:AAA41582.1; PID:q205332

R:Kinet, J.P.; Metzger, H.; Hakimi, J.; Kochan, J.

Biochemistry 26, 4605-4610, 1987

A:Title: A cDNA presumptively coding for the alpha subunit of the receptor with high

A:Reference number: A27116; MUID:88024987

A:Accession: A27116

A:Molecule type: mRNA

A:Residues: 1, 'G', '3-236', 'N', '238-244', 'RLKPNS' <KIN>

R:Tepler, I.; Shimizu, A.; Leder, P.

J. Biol. Chem. 264, 5912-5915, 1989

A:Title: The gene for the rat mast cell high affinity IgE receptor alpha chain. Struc

A:Reference number: I55304; MUID:89174653

A:Accession: I55304

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-16 <RES>

A:Cross-references: GB:M25334; NID:q341335; PIDN:AAA44562.1; PID:q556391

C:Superfamily: Fc gamma receptor III; immunoglobulin homology

C:Keywords: immunoglobulin receptor; transmembrane protein

F:1-23/Domain: signal sequence <status predicted> <SIG>

F:24-245/Product: IgE receptor alpha chain #status predicted <MAT>

F:42-93/Domain: immunoglobulin homology <IMM>

Query Match 79.5%; Score 31; DB 2; Length 245;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
 |||||:
 DB 153 SYDSSNV 159

RESULT 11

G84781

hypothetical protein Atg36540 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84781

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: G84781

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-249 <STO>

A:Cross-references: GB:AE002093; NID:g4581152; PIDN:AA024636.1; GSPDB:GN00139

C:Genetics:

A:Gene: Atg36540

A:Map position: 2

Query Match

79.5%; Score 31; DB 2; Length 249;

Best Local Similarity 71.4%; Pred. No. 35;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
 |||||:
 DB 187 SYDSSNV 193

RESULT 12

T15520

hypothetical protein C15H9.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000

C:Accession: T15520

R:Bentley, D.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid C15H9.

A:Reference number: Z18364

A:Accession: T15520

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-261 <DEN>

A:Cross-references: EMBL:U56965; NID:g1945492; PID:g1293841; PIDN:AA852662.1; GSPDB:GN00

A:Experimental source: strain Bristol N2; clone C15H9

C:Genetics:

A:Gene: CESP:C15H9.2

A:Map position: X

A:Introns: 16/3; 40/2; 78/3; 109/2; 137/3; 192/1; 244/3

Query Match

79.5%; Score 31; DB 2; Length 261;

Best Local Similarity 85.7%; Pred. No. 37;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSSNV 8
 :|||||
 DB 122 YDSSNV 128

RESULT 13

JC4584

insulin-like growth factor binding protein-5 precursor - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 10-Apr-1996 #sequence_revision 24-May-1996 #text_change 05-Nov-1999

C:Accession: JC4584; G23734

R:White, M.E.; Diao, R.; Hachaway, M.R.; Mickelson, J.; Dayton, W.R.

Biochem. Biophys. Res. Commun. 218, 248-253, 1996

A:Title: Molecular cloning and sequence analysis of the porcine insulin-like growth f

A:Reference number: JC4584; MUID:96136309

A:Accession: JC4584

A:Molecule type: mRNA

A:Residues: 1-271 <WHT>

A:Cross-references: GB:U41340; NID:g1173906; PIDN:AA87859.1; PID:g1173907

A:Experimental source: skeletal muscle

R:Shimasaki, S.; Gao, L.; Shimonaka, M.; Ling, N.

Mol. Endocrinol. 5, 938-948, 1991

A:Title: Isolation and molecular cloning of insulin-like growth factor-binding protel

A:Reference number: A23744; MUID:92049376

A:Accession: G23734

A:Molecule type: protein

A:Residues: 20-25, 'X', 27-28, 'X', 30-36, 'X', 38-39 <SHU>

C:Comment: This protein has essential roles in the regulation and coordination of ins

lays a role during myoblast proliferation and differentiation, and is important in th

C:Superfamily: Insulin-like growth factor binding protein 1; thyroglobulin type I rep

C:Keywords: differentiation; growth factor; skeletal muscle

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-271/Product: insulin-like growth factor binding protein-5 #status experimental <

F:191-269/Domain: thyroglobulin type I repeat homology <THY1>

Query Match

79.5%; Score 31; DB 2; Length 271;

Best Local Similarity 85.7%; Pred. No. 38;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
 |||||:
 DB 264 SYDSSNV 270

RESULT 14

T21513

hypothetical protein F28G4.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21513

R:Wall, M.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19433

A:Accession: T21513

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-374 <WIL>

A:Cross-references: EMBL:Z93181; PIDN:CA807605.1; GSPDB:GN00023; CESP:F28G4.2

A:Experimental source: clone F28G4

C:Genetics:

A:Gene: CESP:F28G4.2

A:Map position: 5

A:Introns: 25/3; 49/3; 79/3; 173/3; 233/3; 324/2

Query Match

79.5%; Score 31; DB 2; Length 374;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSSNV 7
 |||||:
 DB 80 YDSSNV 85

RESULT 15

A70795

hypothetical protein RV3712 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: A70795
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98293987
 A:Accession: A70795
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-413 <COL>
 A:Cross-references: GB:AL022121; GB:AL123456; NID:93261559; PIDN:CAI8034.1; PID:e126457
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: RV3712

Query Match 79.5%; Score 31; DB 2; Length 413;
 Best Local Similarity 75.0%; Pred. NO. 60;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 SYSSNV 8
 :||| |||
 Db 174 AYDSPNV 181

Search completed: November 19, 2001, 08:14:24
 Job time: 197 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:36 ; Search time 27.32 Seconds

(without alignments)
10,031 Million cell updates/sec

Title: US-09-610-118-66

Sequence: 1 SYDSNV 8

Scoring table: BLAST62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SWISSPROT_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	84.6	617	VAAL_MOUSE	P50516 mus musculus
2	32	82.1	112	LY6B_HUMAN	P01722 homo sapien
3	32	82.1	460	COXA_BP1P1	O80297 bacterioph
4	32	82.1	469	AMY_CLOAB	P23671 clostridium
5	31	79.5	245	PCCL_RAT	P12371 ratius norv
6	31	79.5	271	IBP5_PIG	Q28985 sus scrofa
7	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
8	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
9	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
10	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
11	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
12	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
13	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
14	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
15	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
16	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
17	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
18	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
19	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
20	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
21	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
22	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
23	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
24	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
25	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
26	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
27	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
28	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
29	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
30	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
31	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
32	31	79.5	272	IBP5_PIG	Q28985 sus scrofa
33	31	79.5	272	IBP5_PIG	Q28985 sus scrofa

34	28	71.8	209	1	RHO1_YEAST	P06780 saccharomyc
35	28	71.8	221	1	EPD2_ONCMY	P28771 oncothychnu
36	28	71.8	261	1	AR5A_PHAVU	O42460 phaseolus v
37	28	71.8	261	1	AR5A_PHAVU	O42460 phaseolus v
38	28	71.8	271	1	IBP5_MOUSE	O07079 mus musculu
39	28	71.8	271	1	IBP5_MOUSE	O07079 mus musculu
40	28	71.8	272	1	IBP5_MOUSE	O07079 mus musculu
41	28	71.8	273	1	IBP5_MOUSE	O07079 mus musculu
42	28	71.8	279	1	IBP5_MOUSE	O07079 mus musculu
43	28	71.8	290	1	IBP5_MOUSE	O07079 mus musculu
44	28	71.8	322	1	IBP5_MOUSE	O07079 mus musculu
45	28	71.8	331	1	IBP5_MOUSE	O07079 mus musculu

ALIGNMENTS

RESULT 1
ID VAAL_MOUSE STANDARD: PRT: 617 AA.
AC P50516;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A, UBIQUITOUS ISOFORM
DE (EC 3.6.1.34) (V-ATPASE A SUBUNIT 1) (VACUOLAR PROTON PUMP ALPHA
DE SUBUNIT 1) (V-ATPASE 69 KDA SUBUNIT 1).
GN ATP6A1 OR ATP6A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96362668; PubMed=8741845;
RA Iatata T., Howell M.L., Dean G.E., Vaananen H.K.;
RT "Resorption-cycle-dependent polarization of mRNAs for different
RT subunits of V-ATPase in bone-resorbing osteoclasts.";
RL Mol. Biol. Cell 7:129-142(1996).
CC -1- FUNCTION: CATALYTIC SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF
CC VACUOLAR ATPASE. V-ATPASE VACUOLAR ATPASE IS RESPONSIBLE FOR
CC ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC
CC CELLS.
CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNIT A, B,
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.ch).
CC
CC EMBL: U13837; AAC52410.1; -
CC MGD: MGI:1201780; At6g2.
DR Interpro: IPR000194; -
DR Interpro: IPR000793; -
DR Pfam: PF00006; ATP-synt_ab; 1.
DR Pfam: PF00306; ATP-synt_ab; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; 1.
KW ATP synthesis; Hydrogen ion transport; Hydrolase; ATP-binding;
KW Multigene family.
FT NP_BIND 250 ATP (POTENTIAL).
SQ
SEQUENCE 617 AA: 68266 MW: 77802843070200CF CRC64;

Query Match 84.6%; Score 33; DB 1; Length 617;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 8
 DB 184 NYDASNV 191

RESULT 2
 LV6B_HUMAN STANDARD; PRT; 112 AA.
 AC P01722;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG LAMBDA CHAIN V-VI REGION NIG-48.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBL_TaxID=9606;
 RX MEDLINE=80094390; PubMed=118171;
 RA Takahashi N., Takayasu T., Isobe T., Shinoda T., Okuyama T.,
 RA Shimizu A.;
 RT "Comparative study on the structure of the light chains of human
 immunoglobulins. II. Assignment of a new subgroup.";
 RL J. Biochem. 86:1523-1535(1979).
 CC -1- MISCELLANEOUS: THIS IS A BENGE-JONES PROTEIN.
 DR HSSP: P01703; 7FAB.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region; Bence-Jones protein.
 FT NON_TER 112
 SQ SEQUENCE 112 AA; 12152 MW; CFB8307BC527A384 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 112;
 Best Local Similarity 85.7%; Pred. No. 3.4;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
 DB 93 SYDSSNV 99

RESULT 3
 COAL_BP1F1 STANDARD; PRT; 460 AA.
 AC O80257;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE COAT PROTEIN A PRECURSOR (ATTACHMENT PROTEIN).
 GN ITI OR 3.
 OS Bacteriophage IF1.
 OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
 NC NCBL_TaxID=10868;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hill D.F., Hughes G., McNaughton J.C., Stockwell P.A., Petersen G.B.;
 RT "DNA sequence of the filamentous coliphage IF1.";
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: COAT PROTEIN A IS NECESSARY FOR ADSORPTION OF THE VIRION
 ONTO THE F-PILUS OF THE HOST CELL (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U02303; AAC62155.1; -;
 KW Phage recognition; Coat protein; Signal.
 FT SIGNAL 1
 FT CHAIN 20 460 COAT PROTEIN A.
 SQ SEQUENCE 460 AA; 48790 MW; 9AD651968C715AB2 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 460;
 Best Local Similarity 85.7%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
 DB 229 SYDSSNV 235

RESULT 4
 AMY_CLOAB STANDARD; PRT; 469 AA.
 AC P23671;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE PUTATIVE ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN
 DE GLUCANHYDROLASE) (FRAGMENT).
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacilli; Clostridium group; Clostridiaceae;
 OC Clostridium.
 NC NCBL_TaxID=1488;
 RX MEDLINE=91072241; PubMed=2254264;
 RC STRAIN=DSM 792;
 RP SEQUENCE FROM N.A.
 RA Gerischer U., Duerte P.;
 RT "Cloning, sequencing, and molecular analysis of the acetate
 RT decarboxylase gene region from Clostridium acetobutylicum.";
 RL J. Bacteriol. 173:6907-6918(1990).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M55392; AAA63759.1; -;
 DR PIR: B37837; B37837.
 KW Hydrolyase; Glycosidase; Carbohydrate metabolism.
 FT NON_TER 1
 SQ SEQUENCE 469 AA; 50870 MW; DB1B670209B01572 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 469;
 Best Local Similarity 85.7%; Pred. No. 17;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSSNV 8
 DB 286 YDSSNV 292

RESULT 5
 FCE1_RAT STANDARD; PRT; 245 AA.
 AC P12371;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR ALPHA-SUBUNIT PRECURSOR
 DE (FCER1) (IG E FC RECEPTOR, ALPHA-SUBUNIT) (FC-EPSILON RI-ALPHA).
 GN FCER1A OR FCER1A.
 OS Rattus norvegicus (Rat).
 OC Charybdis: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88024987; PubMed=2959318;
 RA Kinet J.-P., Metzger H., Hakimi J., Kochan J.;
 RT "A CDNA presumptively coding for the alpha subunit of the receptor
 RT with high affinity for immunoglobulin E.";
 RL Biochemistry 26:4605-4610(1987).
 RN [2]
 RP REVISIONS.
 RA Kochan J.;
 RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=mast cells;
 RX MEDLINE=88158102; PubMed=2964640;
 RA Shimizu A., Tepler I., Benfey P.N., Berenstein E.H., Siraganian R.P.,
 RA Leder P.;
 RT "Human and rat mast cell high-affinity immunoglobulin E receptors:
 RT characterization of putative alpha-chain gene products.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988).
 RN [4]
 RP SEQUENCE OF 21-245 FROM N.A.
 RX MEDLINE=88289772; PubMed=2969594;
 RA Liu F.-T., Albrandt K., Robertson M.W.;
 RT "cDNA heterogeneity suggests structural variants related to the high-
 RT affinity IgE receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:5639-5643(1988).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH
 CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
 CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
 CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
 CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
 CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, AND TWO
 CC DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17153; AAA42045.1;
 DR EMBL: J03606; AAA41582.1;
 DR EMBL: M21622; AAA41146.1;
 DR PIR: A27116; A27116;
 DR PIR: A30154; A30154;
 DR PIR: C31327; C31327;
 DR HSP: P12319; IALT.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 1g. 2.
 DR IGF-binding protein: Receptor: Transmembrane; Glycoprotein; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 23
 FT CHAIN 24 245
 FT TRANSMEM 24 204
 FT TRANSMEM 205 223
 FT DOMAIN 224 245
 FT DOMAIN 42 98
 FT DOMAIN 123 181
 FT CARBOHYD 52 52
 HIGH AFFINITY IMMUNOGLOBULIN EPSILON
 RECEPTOR ALPHA-SUBUNIT.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 IG-LIKE DOMAIN.
 IG-LIKE DOMAIN.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 245 AA; 27793 MW; A0E67D363872197 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 245;
 Best Local Similarity 71.4%; Pred. No. 13;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYDSNV 7
 Db 153 SYDSNMI 159

RESULT 6
 ID IBP5_PIG STANDARD; PRT: 271 AA.
 AC 028985;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 PRECURSOR (IGFBP-5)
 DE (IBP-5) (IGF-BINDING PROTEIN 5).
 GN IGFBP5.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Sketletal muscle;
 RX MEDLINE=96135309; PubMed=8573141;
 RA White M.E., Diao R., Hathaway M.R., Mickelson J., Dayton W.R.;
 RT "Molecular cloning and sequence analysis of the porcine insulin-like
 RT growth factor binding protein-5 complementary deoxyribonucleic
 RT acid.";
 RL Biochem. Biophys. Res. Commun. 218:248-253(1996).
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
 CC PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U41340; AAA87859.1;
 DR InterPro: IPR00716;
 DR InterPro: IPR000867;
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00086; thyroglobulin_1; 1.
 DR PROSITE: PS00222; IGF-BINDING; 1.
 DR PROSITE: PS00484; THYROGLOBULIN_1; 1.
 KW Growth factor binding: Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 271
 FT TRANSMEM 20 271
 FT DOMAIN 214 262
 FT DOMAIN 271 AA; 30323 MW; 95C254E5DC4ED8D CRC64;
 SO SEQUENCE

```

Query Match      79.5%  Score 31; DB 1; Length 271;
Best Local Similarity 85.7%  Pred. NO. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDSSNV 7
1:|||||
DB 264 YDSSNV 270

RESULT 7
SER3_DROME      STANDARD:      PRT:      272 AA.
ID  SER3_DROME      PRT:      272 AA.
AC  P17207; Q9VAB8;
DT  01-AUG-1990 (Rel. 15; Created)
DT  01-OCT-2000 (Rel. 40; Last sequence update)
DT  01-OCT-2000 (Rel. 40; Last annotation update)
DE  SERINE PROTEASE 3 PRECURSOR (EC 3.4.21.-).
GN  SER99DC OR SER3 OR CG17951.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxId=7227;

RN  [1]
RS  SEQUENCE FROM N.A.
RC  STRAIN=BERKELEY;
RX  MEDLINE=20196006; PubMed=10731132;
RA  Adams M.D., Ceolnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sulten G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C.R., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Abrell J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA  Ballif R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA  Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Foster C., Garg J.A., Gary N.S., Gelbart W.M., Glasser K.,
RA  Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Murzy D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson R., Skupski M.P., Smith T.,
RA  Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Szylkas R., Tector C., Turner R., Venner E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster."
RN  [2]
RS  SEQUENCE OF 212-272 FROM N.A.
RC  MEDLINE=89219063; PubMed=2469005;
RX  Yun Y., Davis R.L.;
RT  "Levels of RNA from a family of putative serine protease genes are
RT  reduced in Drosophila melanogaster dunce mutants and are regulated by
RT  cyclic AMP."
RL  Mol. Cell. Biol. 9:692-700(1989).

```

```

CC  -1- FUNCTION: ITS MAJOR FUNCTION MAY BE TO AID IN DIGESTION.
CC  -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THE LARVAL GUT.
CC  -1- DEVELOPMENTAL STAGE: BEGAN TO APPEAR AT LATE EMBRYO STAGE AND
CC  CONTINUED TO INCREASE IN ABUNDANCE THROUGHOUT THE LARVAL STAGE.
CC  THEY ARE NOT PRESENT IN PUPAE BUT REAPPEARED IN THE ADULT.
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC  TRYPSIN FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: AE003771; AAF56973.1; -.
CC  PIR: PS0049; PS0049.
CC  HSSP: P00760; 1A07.
CC  FLYbase: FBgn0003358; Ser99DC.
CC  InterPro: IPR001254; -.
CC  InterPro: IPR001314; -.
CC  Pfam: PF00089; trypsin; 1.
CC  PRINTS: PR00722; CHYMOTRYPSIN.
CC  PROSITE: PS00134; TRYPSIN_HIS; 1.
CC  PROSITE: PS00135; TRYPSIN_SER; 1.
CC  KD  Hydrolysis: Serine protease; Signal; zymogen; Multigene family.
CC  FT  SIGNAL 1 17
CC  FT  PROPEP 18 40
CC  FT  CHAIN 41 272
CC  FT  ACT_SITE 127 127
CC  FT  ACT_SITE 222 222
CC  FT  ACT_SITE 222 222
CC  FT  ACT_SITE 69 85
CC  FT  DISULFID 193 208
CC  FT  DISULFID 218 246
CC  FT  CONFLICT 212 212
CC  FT  CONFLICT 212 212
CC  SO  SEQUENCE 272 AA; 29620 MW; A61E9DE3AFC93AFC CAC64;

Query Match      79.5%  Score 31; DB 1; Length 272;
Best Local Similarity 85.7%  Pred. NO. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNV 8
1:|||||
DB 171 YDSSNV 177

RESULT 8
PRZN_RENSA      STANDARD:      PRT:      548 AA.
ID  PRZN_RENSA      PRT:      548 AA.
AC  P55111;
DT  01-OCT-1996 (Rel. 34; Created)
DT  01-OCT-1996 (Rel. 34; Last sequence update)
DT  01-OCT-1996 (Rel. 40; Last annotation update)
DE  ZINC METALLOPROTEINASE PRECURSOR (EC 3.4.24.-) (HEMOXYLIN).
GN  HLY.
OS  Renibacterium salmoninarum.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Micrococcales; Micrococcaceae; Renibacterium.
OX  NCBI_TaxId=1646;
RN  [1]
RS  SEQUENCE FROM N.A.
RC  STRAIN=MT444;
RX  MEDLINE=95400487; PubMed=7545509;
RA  Grayson T.H., Evenden A.J., Gilpin M.L., Martin K.L., Munn C.B.;
RT  "A gene from renibacterium salmoninarum encoding a product which
RT  shows homology to bacterial zinc-metalloproteases."
RL  Microbiology 141:1331-1341(1995).
CC  -1- FUNCTION: ZINC METALLOPROTEASE WITH HEMOLYTIC PROPERTIES.
CC  HEMOLYTIC ACTIVITY IS OBSERVED FROM 6 TO 37 DEGREES CELSIUS FOR

```

CC MAMMALIAN ERYTHROCYTES.
CC -1- CORFACTOR: BINDS ONE ZINC ION. BINDS FOUR CALCIUM IONS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NOT SECRETED, BUT PROBABLY REMAINS ATTACHED
CC OR ASSOCIATED WITH THE CELL WALL.
CC -1- INDUCTION: EXPRESSION OF THE HEMOLYSIN IS MODULATED BY THE
CC AVAILABILITY OF IRON.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4 (ZINC METALLOPROTEASE);
CC ALSO KNOWN AS THE THERMOLYSIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X76499; CAA54032.1; -
DR HSSP: P05806; IESP.
DR InterPro: IPR000150; -
DR InterPro: IPR001570; -
DR Pfam: PF01447; Peptidase_M4; 1.
DR PRINTS: PRO0730; THERMOLYSIN.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal.
FT SIGNAL 1 28 POTENTIAL.
FT PROPEP 29 ? POTENTIAL.
FT CHAIN ? 548 ZINC METALLOPROTEINASE.
FT METAL 362 362 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 363 363 BY SIMILARITY.
FT METAL 366 366 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 386 386 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 459 459 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 548 AA; 56974 MW; D2285A681953ICBD CRC64;

Query Match 79.5%; Score 31; DB 1; Length 548;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSNNV 8
Db 35 SLDSSNV 42

RESULT 9
CHS2_YEAST
ID CHS2_YEAST STANDARD: PRT; 963 AA.
AC P14180;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CHITIN SYNTHASE 2 (EC 2.4.1.16) (CHITIN-UDP ACETYL-GLUCOSAMINYL
DE TRANSFERASE 2).
GN CHS2 OR YBR038W OR YBR0407.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN SEQUENCE FROM N.A.
RA MEDLINE=90143137; PubMed=2533436;
RA Silverman S.J.;
RT "Similar and different domains of chitin synthases 1 and 2 of S.
RT cerevisiae: two isozymes with distinct functions.";
RL Yeast 5:459-467(1988).
RH [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Andre B., Cziepluch C., Hein C., Jauniaux J.C., Urrestarazu A.,
RA Vissers S.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.

RN [3]
RP MUTAGENESIS.
RX MEDLINE=9529398; PubMed=7775457;
RA Nagaiashi S., Sudoh M., Ono N., Sawada R., Yamaguchi E., Uchida Y.,
RA Mio T., Takagi M., Aisawa M., Yamada-Okabe H.;
RT "Characterization of chitin synthase 2 of Saccharomyces cerevisiae.
RT Implication of two highly conserved domains as possible catalytic
RT sites.";
RL J. Biol. Chem. 270:13961-13967(1995).
CC -1- FUNCTION: ESSENTIAL FOR SEPTUM FORMATION AND CELL DIVISION. CHS2
CC IS REQUIRED FOR MAINTAINING NORMAL CELL MORPHOLOGY.
CC -1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GLUCOSAMINE + [1,4-(N-ACETYL-
CC BETA-D-GLUCOSAMINYL)](N) -> UDP + [1,4-(N-ACETYL-BETA-D-
CC GLUCOSAMINYL)](N+1).
CC -1- ENZYME REGULATION: REQUIRES PROTEOLYTIC ACTIVATION.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23865; AAA34493.1; -
DR EMBL: Z35907; CAA84980.1; -
DR PIR: S45167; S45167.
DR PIR: A50922; A50922.
DR SGD: S0000242; CHS2.
DR InterPro: IPR002923; -
DR Pfam: PF01644; Chitin synth; 1.
KW Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
FT DOMAIN 1 422
FT TRANSSEM 423 443
FT DOMAIN 444 643
FT TRANSSEM 644 664
FT DOMAIN 665 677
FT TRANSSEM 678 698
FT DOMAIN 699 711
FT TRANSSEM 712 732
FT DOMAIN 733 743
FT TRANSSEM 744 764
FT DOMAIN 765 775
FT TRANSSEM 776 796
FT DOMAIN 797 875
FT TRANSSEM 876 896
FT DOMAIN 897 905
FT TRANSSEM 906 926
FT DOMAIN 927 963
FT MUTAGEN 197 197
FT MUTAGEN 312 312
FT MUTAGEN 355 355
FT MUTAGEN 393 393
FT MUTAGEN 441 441
FT MUTAGEN 441 441
FT MUTAGEN 447 447
FT MUTAGEN 490 490
FT MUTAGEN 492 492
FT MUTAGEN 493 493
FT MUTAGEN 494 494
FT MUTAGEN 497 497
FT MUTAGEN 502 502
FT MUTAGEN 505 505
FT MUTAGEN 506 506
FT MUTAGEN 508 508
FT MUTAGEN 509 509
FT MUTAGEN 514 514
FT MUTAGEN 515 515
FT MUTAGEN 516 516
FT MUTAGEN 521 521

1 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
2 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
3 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
4 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
5 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
6 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
7 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
8 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
N->A: 30% DECREASE OF ACTIVITY.
S->A: 20% INCREASE OF ACTIVITY.
D->A: 10% DECREASE OF ACTIVITY.
D->A: 5% DECREASE OF ACTIVITY.
D->A: LOSS OF ACTIVITY.
D->E: LOSS OF ACTIVITY.
N->A: 80% DECREASE OF ACTIVITY.
O->A: 70% DECREASE OF ACTIVITY.
F->A: 80% DECREASE OF ACTIVITY.
E->A: 90% DECREASE OF ACTIVITY.
Y->A: 95% DECREASE OF ACTIVITY.
S->A: 95% DECREASE OF ACTIVITY.
K->A: 90% DECREASE OF ACTIVITY.
E->A: 80% DECREASE OF ACTIVITY.
S->A: 20% INCREASE OF ACTIVITY.
F->A: 60% DECREASE OF ACTIVITY.
F->A: 70% DECREASE OF ACTIVITY.
G->A: 80% DECREASE OF ACTIVITY.
L->A: 80% DECREASE OF ACTIVITY.
P->A: 90% DECREASE OF ACTIVITY.
G->A: 95% DECREASE OF ACTIVITY.
Y->A: LOSS OF ACTIVITY.

```

FT MUTAGEN 522 522 R->A: 60% DECREASE OF ACTIVITY.
FT MUTAGEN 537 537 R->A: NO CHANGE IN ACTIVITY.
FT MUTAGEN 550 550 H->A: 85% DECREASE OF ACTIVITY.
FT MUTAGEN 556 556 H->A: 95% DECREASE OF ACTIVITY.
FT MUTAGEN 559 559 L->A: 95% DECREASE OF ACTIVITY.
FT MUTAGEN 561 561 E->A: LOSS OF ACTIVITY.
FT MUTAGEN 561 561 E->D: 65% DECREASE OF ACTIVITY.
FT MUTAGEN 561 561 E->Q: LOSS OF ACTIVITY.
FT MUTAGEN 562 562 D->A: LOSS OF ACTIVITY.
FT MUTAGEN 562 562 D->E: LOSS OF ACTIVITY.
FT MUTAGEN 562 562 D->N: LOSS OF ACTIVITY.
FT MUTAGEN 563 563 R->A: LOSS OF ACTIVITY.
FT MUTAGEN 563 563 R->K: 94% DECREASE OF ACTIVITY.
FT MUTAGEN 565 565 L->A: 95% DECREASE OF ACTIVITY.
FT MUTAGEN 589 589 T->A: 70% DECREASE OF ACTIVITY.
FT MUTAGEN 592 592 P->A: 70% DECREASE OF ACTIVITY.
FT MUTAGEN 601 601 Q->A: LOSS OF ACTIVITY.
FT MUTAGEN 601 601 Q->N: LOSS OF ACTIVITY.
FT MUTAGEN 602 602 R->A: LOSS OF ACTIVITY.
FT MUTAGEN 602 602 R->K: 95% DECREASE OF ACTIVITY.
FT MUTAGEN 603 603 R->A: LOSS OF ACTIVITY.
FT MUTAGEN 603 603 R->K: 57% DECREASE OF ACTIVITY.
FT MUTAGEN 602 603 RR->KK: LOSS OF ACTIVITY.
FT MUTAGEN 604 604 R->A: LOSS OF ACTIVITY.
FT MUTAGEN 604 604 R->K: LOSS OF ACTIVITY.
FT MUTAGEN 605 605 W->A: LOSS OF ACTIVITY.
FT MUTAGEN 605 605 W->Y: LOSS OF ACTIVITY.
FT MUTAGEN 607 607 N->A: 95% DECREASE OF ACTIVITY.
SQ SEQUENCE 963 AA; 109881 MW; 4D44A287C0B65B5B CRC64;

```

```

Query Match 79.5%; Score 31; DB 1; Length 963;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 SYDSNV 7
   |||||
Db 47 SYDSNV 53

```

```

RESULT 10
CTA2_BACCI STANDARD; PRT; 964 AA.
ID CTA2_BACCI
AC P70873;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CYCLOISOMALTOOLIGOSACCHARIDE GLUCANOTRANSFERASE PRECURSOR (EC 2.4.1.-)
DE (CITASE).
GN CIT.
OS Bacillus circulans;
OC Bacillus; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OC NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U-155;
RA Oguma T., Kurakawa T., Tobe K., Kiteo S., Kobayashi M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PRODUCES CYCLOISOMALTOOLIGOSACCHARIDE FROM DEXTRAN.
CC -1- SIMILARITY: BELONGS TO FAMILY 66 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D88360; BAAL3595.1; -
CC DR Transferase; Glycosyltransferase; Signal.
CC KW SIGNAL
CC AC 1 30 POTENTIAL.

```

```

FT CHAIN 31 964 CYCLOISOMALTOOLIGOSACCHARIDE
FT SEQUENCE 964 AA; 107208 MW; 8849CDE5E2DE9A68 CRC64;
SQ

```

```

OY 1 SYDSNV 6
   |||||
Db 521 SYDSNV 526

```

```

RESULT 11
REGA_CLOAB STANDARD; PRT; 332 AA.
ID REGA_CLOAB
AC Q45831;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE TRANSCRIPTION REGULATORY PROTEIN REGA.
GN REGA OR REPA.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=262;
RX MEDLINE=95291448; PubMed=7539689;
RA Davidson S.P., Santangelo J.D., Reid S.J., Woods D.R.;
RT "A Clostridium acetobutylicum regulator gene (rega) affecting amylase
RT production in Bacillus subtilis."
RL Microbiology 141:989-996(1995).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF AMYLASE PRODUCTION.
CC -1- SIMILARITY: BELONGS TO THE LACTI FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14685; AAA23275.1; -
CC DR HSSP; P15039; 1PRV.
CC DR InterPro; IPR000843; -.
CC DR InterPro; IPR001761; -.
CC DR Pfam; PF00532; Peripla_BP_1like; 1.
CC DR Pfam; PF00356; lact1; 1.
CC DR PRINTS; PR00036; HTHLACT.
CC DR PROSITE; PS00356; HTH_LACTI_FAMILY_1.
CC DR Transcription regulation; DNA-binding.
CC FT DNA_BIND 5 24 H-T-H MOTIF (POTENTIAL).
CC SQ SEQUENCE 332 AA; 37260 MW; 27CBE52FF9BBE007 CRC64;

```

```

Query Match 76.9%; Score 30; DB 1; Length 332;
Best Local Similarity 62.5%; Pred. No. 31;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 SYDSNV 8
   |||||
Db 165 SYDSNV 172

```

```

RESULT 12
YNBH_YEAST STANDARD; PRT; 429 AA.
ID YNBH_YEAST
AC P53729;

```

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOTHETICAL 48.1 KDA PROTEIN IN SECL2-SSK2 INTERGENIC REGION.
 GN YNR029C OR N3360.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Pohl T.M.;
 RL Submitted (May-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: TO P. DENITRIFICANS COBM.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; 271644; CAA96309.1; -
 DR SGD; S0005312; YNR029C
 DR Hypothetical protein, ATP-binding.
 KW NP_BIND 78 85 ATP (POTENTIAL).
 FT SEQUENCE 429 AA: 48121 MW: 58E033E6885EF764 CRC64;

Query Match 76.9%; Score 30; DB 1; Length 429;
 Best Local Similarity 62.5%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYDSSNVV 8
 ||| |:
 Db 36 SYDGNIV 43

RESULT 13
 MB12_ARATH
 ID MB12_ARATH STANDARD; PRT; 462 AA.
 AC Q9SAVO;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MYROSINASE BINDING PROTEIN-LIKE F5F19.10.
 GN F5F19.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaja V.S., Schwartz J.R., Toriumi M., Yu G., Lenz C., Liu S.,
 RA L.J., Kremetskaia I., Luros J., Gonzalez A., Altali H., Ardujo R.,
 RA Buehler E., Conway A.B., Conn L., Dunn P., Hansen N., Hultzer L.,
 RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE JACALIN LECTIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AC006216; AAD12678.1; -

DR HSP: P18670; 1JAC.
 DR InterPro; IPR001229; -
 DR Pfam; PF01419; Jacalin; 3.
 KW Lectin; Repeat; Multigene family.
 SQ SEQUENCE 462 AA: 50167 MW: 0A86DC8AF4B72319 CRC64;

Query Match 76.9%; Score 30; DB 1; Length 462;
 Best Local Similarity 57.1%; Pred. No. 45;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSSNVV 8
 ||| |:
 Db 87 YDSTNII 93

RESULT 14
 VP87_NPVOP
 ID VP87_NPVOP STANDARD; PRT; 624 AA.
 AC P17930;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CAPSID PROTEIN P87.
 GN P87.
 OS Orygia pseudotsugata multicapsid polyhedrosis virus (OPANPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10450;

CC SEQUENCE FROM N.A.
 RN [1]
 RP MEDLINE=90232722; PubMed=2184573;
 RX MEDLINE=90232722; PubMed=2184573;
 RA Mueller R., Pearson M.N., Russell R.L.Q., Rohmann G.F.;
 RT "A capsid-associated protein of the multicapsid nuclear polyhedrosis
 RT virus of Orygia pseudotsugata: genetic location, sequence,
 RT transcriptional mapping, and immunocytochemical characterization";
 RL Virology 176:133-144(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97271300; PubMed=9126251;
 RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
 RA Rohmann G.F.;
 RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
 RT polyhedrosis virus genome";
 RL Virology 229:381-399(1997).

CC -1- SUBCELLULAR LOCATION: CAPSID-ASSOCIATED.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D13959; BAA03061.1; -
 DR EMBL; U75930; AAC59104.1; -
 DR PIR; A34602; VCNV87.
 KW Coat protein; Repeat.
 FT DOMAIN 256 280 2 X 13 AA TANDEM REPEATS, MOTIF A.
 FT REPEAT 256 268 MOTIF A.
 FT REPEAT 269 280 MOTIF A.
 FT DOMAIN 283 324 3 X 11 AA REPEATS, MOTIF B.
 FT REPEAT 283 293 MOTIF B.
 FT REPEAT 296 306 MOTIF B.
 FT REPEAT 314 324 MOTIF B.
 SQ SEQUENCE 624 AA: 70696 MW: 8573D00964C9157B CRC64;

Query Match 76.9%; Score 30; DB 1; Length 624;
 Best Local Similarity 62.5%; Pred. No. 63;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 8
111111:
DB 352 SYDTSSVI 359

RESULT 15
VF12_FOWPV STANDARD; PRT; 630 AA.
ID VF12_FOWPV
AC P36317;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN F12 HOMOLOG.
OS Fowlpox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP MEDLINE-93139784; PubMed-8380837;
RX Ogawa R., Calvert J.G., Yanagida N., Nazerian K.;
RA "Insertional inactivation of a fowlpox virus homologue of the
RT vaccinia virus F12L gene inhibits the release of enveloped virions.";
RL J. Gen. Virol. 74:55-64(1993).
CC -1- FUNCTION: IS INVOLVED IN THE ENVELOPMENT OR RELEASE OF INFECTIOUS
EXTRACELLULAR VIRIONS.
CC -1- SIMILARITY: SOME, TO PROTEIN F12 OF VACCINIA VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sid.ch/announce/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M8588; AAA47187.1; -
DR EMBL: M8587; AAA43820.1; -
DR PIR: J01894; J01894.
DR PIR: S27934; S27934.
KW Transmembrane; Glycoprotein.
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 630 AA; 74098 MW; CD8E1F102032020F CRC64;

Query Match 76.9%; Score 30; DB 1; Length 630;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSSNV 8
111111:
DB 614 YESSNIV 620

Search completed: November 19, 2001, 08:25:37
Job time: 725 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:25:04 : Search time 81.19 Seconds
(without alignments)
13.037 Million cell updates/sec

Title: US-09-610-118-66
Perfect score: 39
Sequence: 1 SYDSNNV 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	680	10 P93658	P93658 brassica na
2	33	84.6	730	10 Q9ZU23	Q9ZU23 arabidopsis
3	33	84.6	934	2 Q9PMF2	Q9PMF2 campylobact
4	33	84.6	1945	5 O96681	O96681 drosophila
5	33	84.6	1945	5 O9V491	O9V491 drosophila
6	32	82.1	506	5 O9N314	O9N314 caenorhabdi
7	32	82.1	521	5 O9VDQ3	O9VDQ3 drosophila
8	32	82.1	527	5 P91310	P91310 caenorhabdi
9	32	82.1	542	5 O9V9K0	O9V9K0 drosophila
10	32	82.1	571	2 O54121	O54121 staphylococ
11	32	82.1	760	2 O9S429	O9S429 clostridium
12	32	82.1	1938	2 P70983	P70983 bacillus sp
13	31	79.5	130	5 O19343	O19343 caenorhabdi
14	31	79.5	131	5 O9Y0F9	O9Y0F9 drosophila
15	31	79.5	192	5 O9V8A2	O9V8A2 drosophila
16	31	79.5	249	10 O9SJQ4	O9SJQ4 arabidopsis
17	31	79.5	261	5 O18030	O18030 caenorhabdi
18	31	79.5	274	5 O17086	O17086 anopheles s
19	31	79.5	374	5 O17852	O17852 caenorhabdi

20	31	79.5	400	2 O08342	O08342 bacillus sp
21	31	79.5	411	2 O69522	O69522 mycobacteri
22	31	79.5	413	2 O69679	O69679 mycobacteri
23	31	79.5	450	5 O17854	O17854 caenorhabdi
24	31	79.5	514	3 O9Y893	O9Y893 candida alb
25	31	79.5	669	14 O9ORV2	O9ORV2 feline call
26	31	79.5	727	10 O9PZ15	O9PZ15 arabidopsis
27	31	79.5	783	10 O04196	O04196 arabidopsis
28	31	79.5	862	3 O13414	O13414 aspergillus
29	31	79.5	907	5 O9Y145	O9Y145 drosophila
30	30	76.9	126	14 O9WSS9	O9WSS9 feline call
31	30	76.9	127	14 O91194	O91194 human papil
32	30	76.9	145	6 P79318	P79318 sus scrofa
33	30	76.9	229	2 O48551	O48551 lactobacill
34	30	76.9	270	10 O9SRS8	O9SRS8 arabidopsis
35	30	76.9	293	5 O9XTE9	O9XTE9 caenorhabdi
36	30	76.9	311	14 O84443	O84443 parametium
37	30	76.9	343	4 O9NR09	O9NR09 homo sapien
38	30	76.9	352	5 O9VAF3	O9VAF3 drosophila
39	30	76.9	354	10 O9SF94	O9SF94 arabidopsis
40	30	76.9	358	5 O44523	O44523 caenorhabdi
41	30	76.9	429	5 O97284	O97284 plasmodium
42	30	76.9	458	5 O09945	O09945 caenorhabdi
43	30	76.9	462	10 O65187	O65187 arabidopsis
44	30	76.9	472	4 O9NOL8	O9NOL8 homo sapien
45	30	76.9	513	5 O9VIF8	O9VIF8 drosophila

ALIGNMENTS

```
RESULT 1
P93658 PRELIMINARY; PRT; 680 AA.
AC P93658;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE JASMONATE INDUCIBLE PROTEIN.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GLOBAL;
RA Geshl N., Brandt A.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: Y11483; CAA72271.1; -.
DR HSSP: P18670; JAC.
DR Mendel: 14659; Brana:2387;14859.
DR InterPro: IPR001229; -.
DR Pfam: PR01419; Jaccalin; 4.
SQ SEQUENCE 680 AA; 69937 MW; 8CD0CF16C30CA1E2 CRC64;

Query Match 84.6%; Score 33; DB 10; Length 680;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSNNV 8
Db 308 YDSNNIT 314

RESULT 2
Q9ZU23 PRELIMINARY; PRT; 730 AA.
AC Q9ZU23;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
```

DE F5F19.6 PROTEIN.
 GN F5F19.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vysotskaya V.S., Schwartz J.R., Tortum M., Yu G., Lenz C., Liu S.,
 RA Li J., Kremetska I., Luoro J., Gonzalez A., Altafi H., Araujo R.,
 RA Buehler E., Conway A.B., Conn L., Dunn P., Hansen N., Hultzer L.,
 RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F5F19 sequence."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006516; AAD12691.1; -
 DR HSSP: P18670; IJAC.
 DR Interpro: IPR001064; -
 DR Interpro: IPR001229; -
 DR Pfam: PF01419; Jacalin; 3.
 DR PROSITE: PS00225; CRYSTALLIN_BETACAMMA; UNKNOWN_1.
 SQ SEQUENCE 730 AA; 73939 MW; E4323AF993B1D95E CRC64;

Query Match 84.6%; Score 33; DB 10; Length 730;
 Best Local Similarity 71.4%; Pred. No. 74;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 SYDSSNV 8
 Db 511 SYDSSNII 517

RESULT 3
 O9PMF2 PRELIMINARY; PRT; 934 AA.
 AC O9PMF2; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PUTATIVE FORMATE DEHYDROGENASE LARGE SUBUNIT (SELENOCYSTEINE
 CONTAINING) (EC 1.2.1.2).
 GN FDHA OR CJ1511.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parrilli J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagsi K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream N.A., Rutherford K.M., Van Vleet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 reveals hypervariable sequences."
 RL Nature 403:665-668(2000).
 DR EMBL: AL139078; CAB7932.1; -
 DR Interpro: IPR001467; -
 DR Pfam: PF00384; molybdopterin; 1.
 DR Pfam: PF01568; molybdop_binding; 1.
 DR PROSITE: PS00551; MOLYBDOPTERIN_PROK_1; UNKNOWN_1.
 SQ SEQUENCE 934 AA; 103906 MW; 5613822028D643DA CRC64;

Query Match 84.6%; Score 33; DB 2; Length 934;
 Best Local Similarity 62.5%; Pred. No. 96;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSSNV 8

Db 714 SYDSSNII 721

RESULT 4
 O96681 PRELIMINARY; PRT; 1945 AA.
 AC O96681; 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PLEXIN A.
 GN PLEXA OR CG11081.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99091049; PubMed=9875845;
 RA Winberg M.L., Noordermeer J.N., Tamagnone L., Comoglio P.M.,
 RA Spriggs M.K., Tessier-Lavigne M., Goodman C.S.;
 RT "Plexin A is a neuronal semaphorin receptor that controls axon
 guidance."
 RL Cell 95:903-916(1998).
 DR EMBL: AF106932; AAD09425.1; -
 DR Flybase: FBgn0025741; plexa.
 DR Interpro: IPR000215; -
 DR Interpro: IPR001627; -
 DR Interpro: IPR002165; -
 DR Interpro: IPR002909; -
 DR Pfam: PF01403; Sema; 2.
 DR Pfam: PF01437; Plexin_repeat; 3.
 DR Pfam: PF01833; TIG; 3.
 DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
 DR SMART: SM00429; IPT; 1.
 SQ SEQUENCE 1945 AA; 218278 MW; 09D55655988DBEF1 CRC64;

Query Match 84.6%; Score 33; DB 5; Length 1945;
 Best Local Similarity 62.5%; Pred. No. 21e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 8
 Db 1266 SYDSSNII 1273

RESULT 5
 O9V491 PRELIMINARY; PRT; 1945 AA.
 AC O9V491; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PLEXA PROTEIN.
 GN PLEXA OR CG11081.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazew R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Abdayani A., An H.-J., Andrews-Pfannkoch C., Balowin D.,
 BA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foutel C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:1185-2195(2000).
 DR EMBL: AE003846; AAF59394.1; -
 DR FlyBase: FBgn0025741; plexa.
 DR InterPro: IPR0000215; -
 DR InterPro: IPR001627; -
 DR InterPro: IPR002165; -
 DR InterPro: IPR002909; -
 DR Pfam: PF01437; Sema; 2.
 DR Pfam: PF01437; Plexin; repeat; 3.
 DR Pfam: PF01833; TIG; 3.
 DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
 DR SMART: SM00423; PSI; 1.
 SQ SEQUENCE 1945 AA; 218356 MW; 4C032CA07C83A745 CRC64;

Query Match 84.6%; Score 33; DB 5; Length 1945;
 Best Local Similarity 62.5%; Pred. No. 2.le+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSNNV 8
 Db 1266 SYESSNI 1273

RESULT 6
 O9N314 PRELIMINARY; PRT; 506 AA.
 AC O9N314;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Y51H7C.F. PROTEIN.
 GN Y51H7C.F.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;

RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC024805; AAF59505.1; -
 SQ SEQUENCE 506 AA; 57932 MW; B079349303AA12DB CRC64;

Query Match 82.1%; Score 32; DB 5; Length 506;
 Best Local Similarity 85.7%; Pred. No. 82;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSNNV 7
 Db 257 SYDSTNV 263

RESULT 7
 O9YD03 PRELIMINARY; PRT; 521 AA.
 AC O9YD03;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CG4936 PROTEIN.
 GN CG4936
 OS *Drosophila melanogaster* (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abdayani A., An H.-J., Andrews-Pfannkoch C., Balowin D.,
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foutel C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
 RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zhang X., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287,2185-2195(2000).
 DR EMBL: AE003728; AAF55737.1; -.
 DR HSSP: P08046; 1A1G.
 DR FlyBase: FBgn0038768; CG4936.
 DR InterPro: IPR000823; -.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00096; zF-C2H2; 5.
 DR PROSITE: PS00290; TG_MHC; UNKNOWN_1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 DR SMART: SM00355; znr_C2H2; 1.
 KW DNA-binding; Metal-binding; Zinc-finger.
 SQ SEQUENCE 521 AA; 58753 MW; 195C4FA4B1830E1F CRC64;

Query Match	82.1%;	Score 32;	DB 5;	Length 521;
Best Local Similarity	62.5%;	Pred. No. 85;		
Matches	5;	Conservative	3;	Mismatches 0;
			Indels	0;
			Gaps	0;

```
QY      1 SYDSSNV 8
         |||::|:|
Db      510 SYDTANIV 517
```

RESULT	B	PRT:	527 AA.
PE1310			
PE1310	PRELIMINARY;		
DT	01-MAY-1997	(TREMblrel. 03,	Created)
DT	01-MAY-1997	(TREMblrel. 03,	Last sequence update)
DT	01-MAY-2000	(TREMblrel. 13,	Last annotation update)
DE	COSMID #476E.		

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Anscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burchon J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Miller L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Lalster N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders R., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*.";
RL Nature 368:32-38(1994).

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Favello A., Maggi L., Chlapelli B.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN {3}
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80838; AAC7112.1; -
DR InterPro: IPR002656; -
DR Pfam: PF01757; DUF33; 1.
SQ .SEQUENCE 527 AA; 60919 MW; 1E459340CE9B6AF2 CRC64;

Query match	82.18;	Score 32;	DB 5;	Length 527;
-------------	--------	-----------	-------	-------------

Oy	1	SYDSSNVV	8
		:	
Db	337	SYDSSNAL	344

RESULT	9		
09V9K0		PRELIMINARY;	PRT; 542 AA.
ID	09V9K0.		
AC	09V9K0.		
DT	01-MAR-2000	(TREMblrel. 13, Created)	
DT	01-MAR-2001	(TREMblrel. 16, Last sequence update)	
DT	01-MAR-2001	(TREMblrel. 16, Last annotation update)	
DE	CG7861	PROTEIN.	
GN	CG7861.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephyridiidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		

Accession	Strain	Sequence from N. A.
RP	SEQUENCE FROM N. A.	
RC	STRAIN-BERKELEY;	
RX	MEDLINE-20196006; PubMed-10731132;	

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amentides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Abiri K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,
RA Avril J.F., Aydayan A., An H.-J., Andrews-Frankkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cayley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Harris A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Hariss N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigwan C.,
RA Jalali M., Kaloustian F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Meruliov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy H., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacllet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reimert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sildenafil K.I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stepleton M., Strong R., Sun E.,
RA Switkes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Welshock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang C., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
DR EMBL: AE003784; AAF57288.2;
DR Flybase: FBgn0033055; CG7661.
DR InterPro: IPR000938;
Pfam: PF01302; CAP_Glyc_1.
SEQUENCE 542 AA: 61580 MW: 6D94685B8EA575B5 CRC64:

Query Match	82.1%;	Score 32;	DB 5;	Length 542;
Best Local Similarity	85.7%;	Pred. No. 88;		

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
:|||||
Db 95 YDSSNV 101

RESULT 10
054121 ID 054121 PRELIMINARY; PRT; 571 AA.
AC 054121:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PEPT.
GN PEPT.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-5.
RX MEDLINE=96035882; PubMed=7556197;
RA Meyer C., Bierbaum G., Heldrich C., Reis M., Suling J.,
RA Iglesias-Wind M., Kemper C., Molitor E., Sahl H.G.;
RT "Nucleotide sequence of the lantibiotic Pept biosynthetic gene cluster
RT and functional analysis of Pept and PeptC. Evidence for a role of Pept
RT in thioether formation.";
RL Eur. J. Biochem. 232:478-489(1995).
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL: Z49865; CAA90021.1; -.
DR HSSP: P13569; INBD.
DR InterPro: IPR001140; -.
DR InterPro: IPR001687; -.
DR InterPro: IPR003439; -.
DR InterPro: IPR003593; -.
DR Pfam: PF00005; ABC_tran; 1.
DR Pfam: PF00664; ABC_membrane; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR SMART: SM00382; AAA; 1.
KM ATP-Binding; Transport.
SQ SEQUENCE 571 AA; 64178 MW; 0AC370046CF86149 CRC64;

Query Match 82.1%; Score 32; DB 2; Length 571;
Best Local Similarity 75.0%; Pred. No. 94;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
:|||||
Db 341 SYDSKNVL 348

RESULT 11
09S429 ID 09S429 PRELIMINARY; PRT; 760 AA.
AC 09S429:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR.
GN AMYP.
OS Clostridium acetobutylicum.
OC plasmid pSO1 megaplasmid.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC824;
RA Sebache F., Cornillot E., Croux C., Soucaille P.;

RT "Molecular characterization of amyp, a pSO1 located gene coding the
RT major alpha-amylase of Clostridium acetobutylicum ATCC824, and its use
RT as a reporter system for strain degeneration.";
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RL EMBL: AF164199; AAD47072.1; -.
DR HSSP: P00691; IBAG.
DR InterPro: IPR000461; -.
DR Pfam: PF00128; alpha-amylase; 1.
DR PRINTS: PR0010; ALPHAMYLASE.
KM Signal; Plasmid.
FT SIGNAL 1 45 POTENTIAL.
FT CHAIN 46 760 ALPHA-AMYLASE.
SQ SEQUENCE 760 AA; 83045 MW; 10E146F40252F6D0 CRC64;

Query Match 82.1%; Score 32; DB 2; Length 760;
Best Local Similarity 85.7%; Pred. No. 1,3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNV 8
:|||||
Db 577 YDSSNV 583

RESULT 12
P70983 ID P70983 PRELIMINARY; PRT; 1938 AA.
AC P70983:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ALKALINE AMYLOPULULANASE.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KSM-1378;
RX MEDLINE=96394539; PubMed=8798645;
RA Hatada Y., Igarashi K., Ozaki K., Ara K., Hitomi J., Kobayashi T.,
RA Kawai S., Katada T., Ito S.;
RT "Amino acid sequence and molecular structure of an alkaline
RT amylopululanase from Bacillus that hydrolyzes alpha-1,4 and alpha-1,6
RT linkages in polysaccharides at different active sites.";
RL J. Biol. Chem. 271:24075-24083(1996).
DR EMBL: D78258; BAA11332.1; -.
DR HSSP: P06278; IVJS.
DR InterPro: IPR000461; -.
DR InterPro: IPR000834; -.
DR InterPro: IPR001899; -.
DR Pfam: PF00128; alpha-amylase; 2.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
SQ SEQUENCE 1938 AA; 214917 MW; A29D3F715EB6A72E CRC64;

Query Match 82.1%; Score 32; DB 2; Length 1938;
Best Local Similarity 62.5%; Pred. No. 3,5e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
:|||||
Db 1669 SYDSSDII 1676

RESULT 13
Q19343 ID Q19343 PRELIMINARY; PRT; 130 AA.
AC Q19343:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)

DE F11C1.2 PROTEIN.
CN F11C1.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Palmer S.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnscough R., Anderson K., Baynes C., Berke M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Cratton M., Dear S., Du Z., Durbin R., Favallo A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Koopra A., Saunders D., Shownkeen R.,
RA Smaildon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.*;
RL Nature 368:32-38(1994).
DR EMBL: 254270; CAA91026.1; -;
SQ SEQUENCE 130 AA; 14913 MW; 0066C8FE171BD294 CRC64;

Query Match 79.5%; Score 31; DB 5; Length 130;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
I:|||||
DB 68 SFDSNV 74

RESULT 14
OY0F9 PRELIMINARY; PRT; 131 AA.
AC OY0F9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ANOXIA UPREGULATED PROTEIN.
GN FAU OR CG6544.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S; TISSUE=HEAD;
RX MEDLINE=99097004; PubMed=9878744;
RA Ma E., Xu T., Haddad G.G.;
RT "Gene regulation by O2 deprivation: an anoxia-regulated novel gene in
RT Drosophila melanogaster.";
RL Brain Res. Mol. Brain Res. 63:217-224(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayr A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Melnikov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relneer K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrlaks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith R.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.*"
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF TISSUE
CC RESPONSIVENESS TO OXYGEN DEPRIVATION.
CC -!- TISSUE SPECIFICITY: CONCENTRATED IN LAMINA NEURONS, FIRST OPTIC
CC LOBE NEURONS AND CORTICAL NEURONS OF CENTRAL BRAIN.
CC -!- INDUCTION: BY ANOXIA.
DR EMBL: AF154418; AAD38397.1; -;
DR EMBL: AE003688; AAF54551.1; -;
DR FLYBase: FBgn0020439; fau.
SQ SEQUENCE 131 AA; 14371 MW; 58FE580BF90570D CRC64;

Query Match 79.5%; Score 31; DB 5; Length 131;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSN 6
I:|||||
DB 59 SYDSSN 64

RESULT 15
OYV8A2 PRELIMINARY; PRT; 192 AA.
AC OYV8A2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG14493.
GN CG14493.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chame M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Adili J.F., Agbayani A., An H.-J., Andrews-Piankocch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003801; AAF57769.1; -;
 DR FlyBase: FBgn0034287; CG14493.
 SQ SEQUENCE 192 AA: 20031 MM: 53D9068F96A87BBA CRC64:

Query Match 79.5%; Score 31; DB 5; Length 192;
 Best Local Similarity 62.5%; Pred. No. 48;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSNTV 8
 |:|:|:|:
 Db 78 SFDSTNVI 85

Search completed: November 19, 2001, 08:25:06
 Job time: 744 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:12:36 ; Search time 77.92 Seconds
(without alignments)
6.224 Million cell updates/sec

Title: US-09-610-118-66
Perfect score: 39
Sequence: 1 SYDSSNVV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	39	100.0	8 22	AA61298 Anti-TANGO 268 scf
2	35	89.7	402 15	AA60578 Osteogenic protein
3	33	84.6	730 21	AA629582 Arabidopsis thalia
4	33	84.6	752 21	AA629581 Arabidopsis thalia
5	32	82.1	798 20	AAW73552 Pullulanase protei
6	32	82.1	893 17	AAW09257 Bacillus alkaline
7	32	82.1	1938 17	AAW09255 Bacillus alkaline
8	32	82.1	1938 17	AAW73553 Full length Pullul
9	31	79.5	220 14	AA642340 Subunit of the hum
10	31	79.5	245 20	AA690386 Alpha subunit of r
11	31	79.5	325 20	AA602465 Polypeptide isolat

12	31	79.5	469 12	AA615510 Tomato ACC synthas
13	31	79.5	964 13	AAW32619 Cyclic-Isomaltolig
14	30	76.9	29 18	AA622106 Peptide 3-29 deriv
15	30	76.9	126 30	AAW64528 Papillomavirus maj
16	30	76.9	308 21	AA649677 Arabidopsis thalia
17	30	76.9	380 21	AA635707 Arabidopsis thalia
18	30	76.9	390 21	AA649676 Arabidopsis thalia
19	30	76.9	462 21	AA635706 Arabidopsis thalia
20	30	76.9	618 18	AAW55098 Streptococcus pneu
21	30	76.9	627 18	AAW17885 Photorhabdus lumin
22	30	76.9	627 19	AAW56544 Toxin Tcabi, encod
23	30	76.9	676 18	AAW35853 Human CD18, for use
24	30	76.9	699 20	AAW81840 Human LFA-1 beta c
25	30	76.9	769 9	AA680836 Beta subunit of hu
26	30	76.9	769 11	AA607113 Recombinant beta-s
27	30	76.9	769 13	AA624256 Beta-subunit CD18
28	30	76.9	769 16	AA680108 LFA-1 beta subunit
29	30	76.9	769 21	AA603974 LFA-1 CD18 polypep
30	30	76.9	1189 18	AAW17884 Photorhabdus lumin
31	30	76.9	1189 19	AAW56543 Toxin Tcabi, encode
32	29	74.4	16 20	AAW33192 E. cloacae adhesiv
33	29	74.4	40 18	AAW12037 Curvularia verruc
34	29	74.4	108 17	AA688719 Human antibody lam
35	29	74.4	129 21	AA616381 Pinus radiata flav
36	29	74.4	177 29	AA632755 Eucalyptus grandis
37	29	74.4	246 18	AAW09434 Anti-CD19 antibody
38	29	74.4	247 18	AAW09443 Modified single ch
39	29	74.4	281 19	AAW62314 Human scFvCD19 pro
40	29	74.4	285 20	AAW14200 Mouse DNA demethyl
41	29	74.4	288 19	AAW62316 Mouse OKT3 variant
42	29	74.4	288 19	AAW62317 Mouse bispecific a
43	29	74.4	291 20	AAW14198 Human DNA demethyl
44	29	74.4	469 22	AAW59726 Tomato ACC synthas
45	29	74.4	483 19	AAW82315 Chimeric CD19/CTCR

ALIGNMENTS

RESULT 1
ID AAB61298 standard; Peptide: 8 AA.
XX AAB61298;
XX
XX
XX 04-APR-2001 (first entry)
XX
XX
XX Anti-TANGO 268 scfv CDR, SEQ ID NO: 66.
DE
XX
XX Human: antibody; scfv; CDR; complementarity determining region;
KM TANGO 268; cardiant; cerebroprotective; cyostatic; anticoagulant;
KM thrombolytic; antiarteriosclerotic; haemostatic; glycoprotein VI; GPVI;
KM platelet membrane glycoprotein receptor; bleeding disorder;
KM blood vessel injury; thrombotic disorder; haemorrhagic disorder; stroke;
KM ischaemia; cardiovascular disease; immunological disease; liver disorder;
KM cancer.
XX
XX Homo sapiens.
OS
XX WO200100810-A1.
PN
XX
XX 04-JAN-2001.
PD
XX
XX 30-JUN-2000; 2000MO-US18152.
PF
XX
XX 30-JUN-1999; 99US-0345468.
PR 06-DEC-1999; 99US-0454824.
PR 14-FEB-2000; 2000US-0503387.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Bustfield SJ, Villegas J, Jandrot-Perrus M, Vainchenker W, Gill DS;
PI Qian MD, Kingsbury G;

XX WPI: 2001-080877/09.
DR
XX
PT New genes encoding human platelet-expressed collagen receptor,
PT glycoprotein VI, and its modulators, useful for preventing, treating
PT and diagnosing hemorrhagic disorders, thrombotic diseases and
PT immunological disorders.
XX
XX
PS Claim 32: Page 102; 227pp; English.
XX
XX
CC The present sequence is given in a specification relating to an isolated
CC nucleic acid molecule encoding a platelet membrane glycoprotein receptor
CC glycoprotein VI (GPVI), also called TAMGO 268. The GPVI polynucleotides
CC and polypeptides and their modulators, e.g. antisense nucleic acids,
CC ribozymes and antibodies, are useful for preventing, treating and
CC diagnosing disorders associated with aberrant expression or activity of
CC GPVI. These disorders include bleeding disorders
CC (e.g. thrombocytopaenia), blood vessel injury, thrombotic disorders
CC (e.g. thrombotic occlusion of the coronary arteries), haemorrhagic
CC disorders, coronary artery and cerebral artery diseases (e.g. stroke and
CC ischaemia), cardiovascular diseases (e.g. atherosclerosis and myocardial
CC infarction), immunological diseases (e.g. platelet disorder) and
CC embryonic liver disorders. Preferably they are used to prevent acute
CC cardiac ischaemia following angioplasty and metastatic cancers,
CC especially of the colon and liver.
XX
XX
SQ Sequence 8 AA:

Query Match 100.0%; Score 39; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSNNV 8
| | | | | | | |
Db 1 sydsnnv 8

RESULT 2
AAR60578
ID AAR60578 standard; Protein; 402 AA.
XX
AC AAR60578;
XX
DT 30-MAR-1995 (first entry)
XX
DE Osteogenic protein OP2.
XX
KM Osteogenic protein.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Cleavage-site 239..242
FT /note= "proteolytic cleavage site"
FT Cleavage-site 260..263
FT /note= "proteolytic cleavage site"
XX
XX
PN MO9420539-A.
XX
XX
PD 15-SEP-1994.
XX
PF 04-MAR-1994; 94WO-US02335.
XX
PR 04-MAR-1993; 93US-0027070.
XX
XX
PA (CREA-) CREATIVE BIOMOLECULES INC.
XX
PI Jones WK, Oppermann H, Ozkaynak E, Rueger DC, Sampath KT;
XX 'Tucker RF;
XX WPI: 1994-302971/37.
DR N-PSDB; AA071426.

XX Binding partners, esp. antibodies, specific for different forms
PT of osteogenic protein - for differentiating between mature and
PT soluble complexed forms of the protein in culture media or serum.
PS Disclosure; Page 50-52; 70pp; English.
XX
XX
CC The osteogenic protein is produced recombinantly in mammalian cell
CC cultures, and may be provided to a site for bone induction in a
CC mammal with a suitable matrix to allow infiltration, proliferation
CC and differentiation of migrating progenitor cells.
XX
XX
SQ Sequence 402 AA:

Query Match 89.7%; Score 35; DB 15; Length 402;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSNNV 8
| | | | | | | |
Db 360 ydsnnv 366

RESULT 3
AAG29582
ID AAG29582 standard; Protein; 730 AA.
XX
AC AAG29582;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35222.
XX
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 23-APR-1999; 99US-0130891.
XX
PR 28-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 04-MAY-1999; 99US-0132407.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 07-MAY-1999; 99US-0132487.
XX
PR 11-MAY-1999; 99US-0132863.
XX
PR 14-MAY-1999; 99US-0134256.
XX
PR 14-MAY-1999; 99US-0134218.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999;	990S-0134768.	PR 27-JUL-1999;	990S-0145919.
PR 19-MAY-1999;	990S-0134941.	PR 28-JUL-1999;	990S-0145951.
PR 20-MAY-1999;	990S-0135124.	PR 02-AUG-1999;	990S-0146386.
PR 21-MAY-1999;	990S-0135353.	PR 02-AUG-1999;	990S-0146388.
PR 24-MAY-1999;	990S-0135629.	PR 02-AUG-1999;	990S-0146389.
PR 25-MAY-1999;	990S-0136021.	PR 03-AUG-1999;	990S-0147038.
PR 27-MAY-1999;	990S-0136392.	PR 04-AUG-1999;	990S-0147204.
PR 28-MAY-1999;	990S-0136782.	PR 04-AUG-1999;	990S-0147302.
PR 01-JUN-1999;	990S-0137222.	PR 05-AUG-1999;	990S-0147492.
PR 03-JUN-1999;	990S-0137528.	PR 05-AUG-1999;	990S-0147260.
PR 04-JUN-1999;	990S-0137502.	PR 06-AUG-1999;	990S-0147303.
PR 07-JUN-1999;	990S-0137724.	PR 06-AUG-1999;	990S-0147416.
PR 08-JUN-1999;	990S-0138094.	PR 09-AUG-1999;	990S-0147493.
PR 10-JUN-1999;	990S-0138540.	PR 09-AUG-1999;	990S-0147935.
PR 10-JUN-1999;	990S-0138847.	PR 10-AUG-1999;	990S-0148171.
PR 14-JUN-1999;	990S-0139119.	PR 11-AUG-1999;	990S-0148319.
PR 16-JUN-1999;	990S-0139452.	PR 12-AUG-1999;	990S-0148341.
PR 16-JUN-1999;	990S-0139453.	PR 13-AUG-1999;	990S-0148565.
PR 17-JUN-1999;	990S-0139492.	PR 13-AUG-1999;	990S-0148684.
PR 18-JUN-1999;	990S-0139454.	PR 16-AUG-1999;	990S-0149368.
PR 18-JUN-1999;	990S-0139455.	PR 17-AUG-1999;	990S-0149175.
PR 18-JUN-1999;	990S-0139456.	PR 18-AUG-1999;	990S-0149426.
PR 18-JUN-1999;	990S-0139457.	PR 20-AUG-1999;	990S-0149722.
PR 18-JUN-1999;	990S-0139458.	PR 20-AUG-1999;	990S-0149723.
PR 18-JUN-1999;	990S-0139459.	PR 20-AUG-1999;	990S-0149929.
PR 18-JUN-1999;	990S-0139459.	PR 23-AUG-1999;	990S-0149902.
PR 18-JUN-1999;	990S-0139460.	PR 23-AUG-1999;	990S-0149930.
PR 18-JUN-1999;	990S-0139461.	PR 25-AUG-1999;	990S-0150566.
PR 18-JUN-1999;	990S-0139462.	PR 26-AUG-1999;	990S-0150884.
PR 18-JUN-1999;	990S-0139463.	PR 27-AUG-1999;	990S-0151065.
PR 18-JUN-1999;	990S-0139750.	PR 27-AUG-1999;	990S-0151066.
PR 18-JUN-1999;	990S-0139763.	PR 27-AUG-1999;	990S-0151080.
PR 21-JUN-1999;	990S-0139817.	PR 30-AUG-1999;	990S-0151303.
PR 22-JUN-1999;	990S-0139899.	PR 31-AUG-1999;	990S-0151438.
PR 23-JUN-1999;	990S-0140353.	PR 01-SEP-1999;	990S-0151930.
PR 23-JUN-1999;	990S-0140354.	PR 07-SEP-1999;	990S-0152363.
PR 24-JUN-1999;	990S-0140695.	PR 10-SEP-1999;	990S-0153070.
PR 28-JUN-1999;	990S-0140823.	PR 13-SEP-1999;	990S-0153758.
PR 29-JUN-1999;	990S-0140991.	PR 15-SEP-1999;	990S-0154018.
PR 30-JUN-1999;	990S-0141287.	PR 16-SEP-1999;	990S-0154039.
PR 01-JUL-1999;	990S-0141842.	PR 20-SEP-1999;	990S-0154779.
PR 01-JUL-1999;	990S-0142154.	PR 22-SEP-1999;	990S-0155139.
PR 02-JUL-1999;	990S-0142055.	PR 23-SEP-1999;	990S-0155486.
PR 06-JUL-1999;	990S-0142390.	PR 24-SEP-1999;	990S-0155659.
PR 08-JUL-1999;	990S-0142803.	PR 28-SEP-1999;	990S-0156458.
PR 09-JUL-1999;	990S-0142920.	PR 29-SEP-1999;	990S-0156596.
PR 12-JUL-1999;	990S-0142977.	PR 04-OCT-1999;	990S-0157117.
PR 13-JUL-1999;	990S-0143542.	PR 05-OCT-1999;	990S-0157753.
PR 14-JUL-1999;	990S-0143622.	PR 06-OCT-1999;	990S-0157865.
PR 15-JUL-1999;	990S-0144005.	PR 07-OCT-1999;	990S-0158029.
PR 16-JUL-1999;	990S-0144085.	PR 08-OCT-1999;	990S-0158232.
PR 16-JUL-1999;	990S-0144086.	PR 12-OCT-1999;	990S-0158369.
PR 19-JUL-1999;	990S-0144325.	PR 13-OCT-1999;	990S-0159294.
PR 19-JUL-1999;	990S-0144331.	PR 13-OCT-1999;	990S-0159295.
PR 19-JUL-1999;	990S-0144332.	PR 14-OCT-1999;	990S-0159329.
PR 19-JUL-1999;	990S-0144333.	PR 14-OCT-1999;	990S-0159330.
PR 19-JUL-1999;	990S-0144335.	PR 14-OCT-1999;	990S-0159331.
PR 20-JUL-1999;	990S-0144352.	PR 14-OCT-1999;	990S-0159637.
PR 20-JUL-1999;	990S-0144632.	PR 14-OCT-1999;	990S-0159638.
PR 20-JUL-1999;	990S-0144684.	PR 18-OCT-1999;	990S-0159584.
PR 21-JUL-1999;	990S-0144814.	PR 21-OCT-1999;	990S-0160741.
PR 21-JUL-1999;	990S-0145086.	PR 21-OCT-1999;	990S-0160761.
PR 21-JUL-1999;	990S-0145088.	PR 21-OCT-1999;	990S-0160766.
PR 22-JUL-1999;	990S-0145085.	PR 21-OCT-1999;	990S-0160770.
PR 22-JUL-1999;	990S-0145087.	PR 21-OCT-1999;	990S-0160814.
PR 22-JUL-1999;	990S-0145089.	PR 21-OCT-1999;	990S-0160815.
PR 22-JUL-1999;	990S-0145192.	PR 22-OCT-1999;	990S-0160980.
PR 23-JUL-1999;	990S-0145145.	PR 22-OCT-1999;	990S-0160981.
PR 23-JUL-1999;	990S-0145224.	PR 22-OCT-1999;	990S-0160989.
PR 26-JUL-1999;	990S-0145276.	PR 25-OCT-1999;	990S-0161404.
PR 27-JUL-1999;	990S-0145913.	PR 25-OCT-1999;	990S-0161405.
PR 27-JUL-1999;	990S-0145918.	PR 25-OCT-1999;	990S-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 84.6%; Score 33; DB 21; Length 730;
Best Local Similarity 71.4%; Pred NO. 75;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDSSNVV 8
DB 511 ydesnll 517

RESULT 4
ID AAG29581 standard; Protein: 752 AA.
XX AAG29581;
AC AAG29581;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35221.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
XX EPI033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0122548.
PR 23-MAR-1999; 99US-0122788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

```

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150864.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      84.6%; Score 33; DB 21; Length 752;
Best Local Similarity 71.4%; Pred. No. 78;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDSSNVV 8
    |||||:
Db 533 ydssnll 539

RESULT 5
AAW73552
ID AAW73552 standard; protein; 798 AA.
XX
AC AAW73552;
XX
DT 08-MAR-1999 (first entry)
XX
DE Pullulanase protein.
XX
KW Pullulanase; variant; bleach-containing detergent; oxidation resistant.
XX
OS Bacillus sp.
XX
PN JP10327868-A.
XX
PD 15-DEC-1998.
XX
PE 30-MAY-1997; 97JP-0141596.
XX
PR 30-MAY-1997; 97JP-0141596.
XX
PA (KAOS ) KAO CORP.
XX
DR WPI; 1999-093031/09.
XX
XX New mutant pullulanase - useful in bleach-containing detergents
PT
XX
PS Claim 11; Page 8-10; 19pp; Japanese.
XX
XX
CC This sequence represents the Bacillus pullulanase protein. The invention
CC relates to a mutant pullulanase in which at least one methionine residue
CC in the pullulanase is deleted or replaced by an optional amino acid
CC residue other than Cys and Met and/or at least one alanine residue in the
CC amino acid residues constituting pullulanase is deleted or replaced by an
CC optional amino acid residue other than Ala. The mutant pullulanase is
CC useful in bleach-containing detergents. The pullulanase is highly
CC resistant to oxidation.
CC
XX
SQ Sequence 798 AA;

```

```

Query Match      82.1%; Score 32; DB 20; Length 798;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNVV 8
    |||||:
Db 647 sydssnll 654

RESULT 6
AAW09257
ID AAW09257 standard; Protein; 893 AA.
XX
AC AAW09257;
XX
DT 27-MAR-1997 (first entry)
XX
DE Bacillus alkaline pullulanase.
XX
KW Amylopullulanase; alpha-amylase; pullulanase; detergent additive.
XX
XX

```

```

OS  Bacillus sp. KSM-API378 (FERM BP-3048).
XX
FH  Key
FT  Active-site
FT  Location/Qualifiers
FT  351..467
FT  /note="pullulanase active centre"
FT
FT  Region
FT  351..356
FT  /label="Region-I"
FT  /note="pullulanase active centre region I"
FT
FT  Region
FT  415..423
FT  /label="Region-II"
FT  /note="pullulanase active centre region II"
FT
FT  Region
FT  448..451
FT  /label="Region-III"
FT  /note="pullulanase active centre region III"
FT
FT  Region
FT  531..536
FT  /label="Region-IV"
FT  /note="pullulanase active centre region IV"
FT
XX  WO9635794-A1.
XX  14-NOV-1996.
XX
XX  10-MAY-1996; 96WO-JP01243.
XX
XX  10-MAY-1995; 95JP-0111547.
XX
XX  (KAOS ) KAO CORP.
XX
XX  Ara K, Hatada Y, Igarashi K, Ito S, Kawai S, Ozaki K;
XX  WPI; 1996-518682/51.
XX  N-PSDB; AAT47878.
XX
XX  DNA encoding alkaline pullulanase having alkaline alpha-amylase
XX  activity - useful for improving the efficacy of dish washing
XX  detergents and detergents for clothes, esp. on starch soils
XX
XX  Claim 7; Page 59-62; 75pp; English.
XX
XX  Recombinant alkaline pullulanase (AAW09257) and alkaline alpha-
XX  amylase (AAW09256) can be produced in transformed host cells using
XX  vectors incorporating appropriate fragments of the alkaline
XX  amylopullulanase gene (see also AAT47878) of Bacillus sp. isolate
XX  KSM-API378. The enzymes respectively comprise the N-terminal and
XX  C-terminal moieties of the amylopullulanase. The efficacy of
XX  dish-washing detergents or detergents for clothes is greatly
XX  improved, partic. on starch soils, when the alpha-amylase and
XX  pullulanase are both incorporated into the detergent compsn.
XX
XX  Sequence 893 AA;
XX
XX  Query Match
XX  Best Local Similarity 82.1%; Score 32; DB 17; Length 893;
XX  Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX  1 SYDSSNVV 8
XX  624 sydsed11 631
XX
XX  RESULT 7
XX  AAW09255
XX  ID AAW09255 standard; Protein: 1938 AA.
XX
XX  AAW09255;
XX
XX  27-MAR-1997 (first entry)
XX
XX  Bacillus alkaline amylopullulanase.
XX
XX  Amylopullulanase; alpha-amylase; pullulanase; detergent additive.
XX

```

```

OS  Bacillus sp. KSM-API378 (FERM BP-3048).
XX
XX  Key
XX  Location/Qualifiers
XX  1..32
XX  /label="Sig_peptide"
XX
XX  Protein
XX  33..1938
XX  /label="Mat_protein"
XX  /note="mature amylopullulanase"
XX
XX  Domain
XX  33..833
XX  /label="Alpha-amylase"
XX  /note="N-terminal alpha-amylase"
XX
XX  Active-site
XX  462..645
XX  /note="amylase active centre"
XX
XX  Region
XX  462..467
XX  /label="Region-I"
XX  /note="amylase active centre region I"
XX
XX  Region
XX  546..554
XX  /label="Region-II"
XX  /note="amylase active centre region II"
XX
XX  Region
XX  579..582
XX  /label="Region-III"
XX  /note="amylase active centre region III"
XX
XX  Region
XX  640..645
XX  /label="Region-IV"
XX  /note="amylase active centre region IV"
XX
XX  Region
XX  834..866
XX  /note="intervening sequence"
XX
XX  Region
XX  944..976
XX  /note="intervening sequence"
XX
XX  Domain
XX  1046..1938
XX  /label="Pullulanase"
XX  /note="C-terminal alkaline pullulanase"
XX
XX  Active-site
XX  1396..1481
XX  /note="pullulanase active centre"
XX
XX  Region
XX  1396..1401
XX  /label="Region-I"
XX  /note="pullulanase active centre region I"
XX
XX  Region
XX  1460..1468
XX  /label="Region-II"
XX  /note="pullulanase active centre region II"
XX
XX  Region
XX  1493..1496
XX  /label="Region-III"
XX  /note="pullulanase active centre region III"
XX
XX  Region
XX  1576..1581
XX  /label="Region-IV"
XX  /note="pullulanase active centre region IV"
XX
XX  WO9635794-A1.
XX  14-NOV-1996.
XX
XX  10-MAY-1996; 96WO-JP01243.
XX
XX  10-MAY-1995; 95JP-0111547.
XX
XX  (KAOS ) KAO CORP.
XX
XX  Ara K, Hatada Y, Igarashi K, Ito S, Kawai S, Ozaki K;
XX  WPI; 1996-518682/51.
XX  N-PSDB; AAT47878.
XX
XX  DNA encoding alkaline pullulanase having alkaline alpha-amylase
XX  activity - useful for improving the efficacy of dish washing
XX  detergents and detergents for clothes, esp. on starch soils
XX
XX  Disclosure; Page 42-54; 75pp; English.
XX
XX  Amylopullulanase (AAW09255) of Bacillus sp. isolate KSM-API378 has an
XX  N-terminal alkaline alpha-amylase moiety (see also AAW09256) and a
XX  C-terminal alkaline pullulanase moiety (see also AAW09257). It is
XX  the product of an open reading frame in a DNA fragment (AAT47878)
XX  isolated using shotgun cloning and PCR (see also AAT47879-84). The

```

CC amylopullulanase, alpha-amylase and pullulanase can each be
 CC expressed in transformed host cells using vectors contg. the
 CC appropriate gene fragments. The enzymes are useful as additives
 CC in dish-washing and clothes detergent compsns.
 XX

SQ Sequence 1938 AA;

Query Match 82.1%; Score 32; DB 17; Length 1938;
 Best Local Similarity 62.5%; Pred. No. 3.8e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 8
 |||||
 Db 1669 sydsdhl 1676

RESULT 8

ID AAW73553 standard; protein; 1938 AA.

AC AAW73553;

DT 08-MAR-1999 (first entry)

DE Full length Pullulanase protein.

XX Pullulanase; variant; bleach-containing detergent; oxidation resistant.

OS Bacillus sp.

Key Location/Qualifiers

FT Misc-difference 121..137
 /note= "nucleotides encoding these residues are not present in the coding sequence given in the specification"

FT Misc-difference 1050..1065
 /note= "nucleotides encoding these residues are not present in the coding sequence given in the specification"

FT Misc-difference 1050..1065
 /note= "nucleotides encoding these residues are not present in the coding sequence given in the specification"

FT Misc-difference 1050..1065
 /note= "nucleotides encoding these residues are not present in the coding sequence given in the specification"

PN JP10327868-A.

XX 15-DEC-1998.

PF 30-MAY-1997; 97JP-0141596.

PR 30-MAY-1997; 97JP-0141596.

PA (KAOS) KAO CORP.

WI 1999-099031/09.

DR N-PSDB; AAV62881.

PT New mutant pullulanase - useful in bleach-containing detergents

PS Claim 13; Page 10-17; 19pp; Japanese.

CC This sequence represents the Bacillus pullulanase protein. The invention
 CC relates to a mutant pullulanase in which at least one methionine residue
 CC in the pullulanase is deleted or replaced by an optional amino acid
 CC residue other than Cys and Met and/or at least one alanine residue in the
 CC amino acid residues constituting pullulanase is deleted or replaced by an
 CC optional amino acid residue other than Ala. The mutant pullulanase is
 CC useful in bleach-containing detergents. The pullulanase is highly
 CC resistant to oxidation.

SQ Sequence 1938 AA;

Query Match 82.1%; Score 32; DB 20; Length 1938;
 Best Local Similarity 62.5%; Pred. No. 3.8e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 8
 |||||
 Db 1669 sydsdhl 1676

RESULT 9

ID AAR42340 standard; Protein; 220 AA.

AC AAR42340;

DT 21-JUN-1994 (first entry)

DE Subunit of the human IGE receptor.

XX IGE; immunoglobulin E receptor; beta subunit; basophils; allergy;

KW aggregation; signal transduction; diagnosis; antagonist.

OS Homo sapiens.

PN WO9321317-A.

PD 28-OCT-1993.

PF 16-APR-1993; 93WO-US03419.

PR 16-APR-1992; 92US-0869933.

PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

PI Kinet JP;

DR WPI; 1993-351727/44.

XX Immunoglobulin E receptor human beta sub-unit isolation - using

PT 1st strand reverse transcripts from human basophils as templates

PT for a polymerase chain reaction, used to treat and diagnose

PS allergic diseases

XX Disclosure; Fig 12; 154pp; English.

XX The sequence is that of a subunit of the human IGE receptor.

CC The protein can be used to identify human beta subunit

CC FcεRI inhibitors (immunoglobulin E receptor) which inhibit the

CC binding of IGE to its receptor and inhibit the aggregation function

CC of the receptor or the signal transducing function related to

CC allergic response. Such inhibitors can be used for the treatment or

CC prevention of allergic disease.

CC See also AAR42337-42.

XX See also AAR42337-42.

SQ Sequence 220 AA;

Query Match 79.5%; Score 31; DB 14; Length 220;
 Best Local Similarity 71.4%; Pred. No. 51;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
 |||||
 Db 130 sydsnml 136

RESULT 10

ID AAP90386 standard; protein; 245 AA.

AC AAP90386;

DT 01-NOV-1989 (first entry)

DE Alpha subunit of rat mast cell IGE surface receptor.

```

KW Immunoglobulin E receptor alpha subunit;
KM allergies; non-peptide drug design; rat.
XX Rat.
OS
XX WO8905352-A.
XX 15-JUN-1989.
XX
XX 29-NOV-1988; 88WO-US04255.
XX
XX 01-DEC-1987; 87US-0127214.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Streganlian R, Shimizu A, Leder P, Benfey P;
XX
XX WPI; 1989-192698/26.
XX
XX N-PSDB; AAN90125.
XX
XX CDNA encoding IGE receptor alpha-subunit - used to treat allergies.
XX
XX Disclosure: fig 4; 17pp; English.
XX
XX Immunoglobulin E receptor alpha subunit of rat mast cell IGE
XX surface receptor (see corresp. AAN90125) . Used to produce antibodies
XX which can diagnose IGE receptor levels, measure and treat allergies,
XX and design non-peptide drugs.
XX
XX Sequence 245 AA:

Query Match          79.5%; Score 31; DB 10; Length 245;
Best Local Similarity 71.4%; Pred. No. 57;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
Db 153 sydsnm1 159

RESULT 11
AAY02465
ID AAY02465 standard; protein: 325 AA.
XX
XX AAY02465;
XX
XX 13-JUL-1999 (first entry)
XX
XX Polypeptide isolated from prostate tissue.
XX
XX N-terminal; prostate tumour cell; immunogenic; treatment; diagnosis;
XX prostate cancer.
XX
XX Homo sapiens.
XX
XX WO9918210-A2.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-US21166.
XX
XX 23-JUN-1998; 98US-0102679.
XX
XX 07-OCT-1997; 97US-0946026.
XX
XX (CORI-) CORIAX CORP.
XX
XX Dillon DC, Mitcham JL, Reed SG, Twardzik DR;
XX
XX WPI; 1999-277272/23.
XX
XX New isolated prostate polypeptides useful for the treatment,
XX diagnosis and monitoring of prostate cancer
PT

```

```

XX
XX Example 6; Page 93-94; 106pp; English.
XX
XX The present sequence represents a polypeptide obtained from prostate
XX tumour cells. The polypeptide comprises an immunogenic portion of a
XX prostate protein. The polypeptides and DNA obtained from prostate
XX tumour cells, as well as antibodies raised against the protein, can
XX be used in the treatment, diagnosis and monitoring of prostate cancer.
XX
XX Sequence 325 AA:

Query Match          79.5%; Score 31; DB 20; Length 325;
Best Local Similarity 85.7%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 7
Db 97 syessnv 103

RESULT 12
AAR15510
ID AAR15510 standard; protein: 469 AA.
XX
XX AAR15510;
XX
XX 09-MAR-1992 (first entry)
XX
XX Tomato ACC synthase encoded by clone LE-ACC1.
XX
XX 1-aminocyclopropane-1-carboxylic acid synthase; fruit ripening.
XX
XX Lycopersicon esculentum c.v. Rutgers.
XX
XX US7579896-A.
XX
XX 12-NOV-1991.
XX
XX 10-SEP-1990; 90US-0579896.
XX
XX 10-SEP-1990; 90US-0579896.
XX
XX (USDA ) US SEC OF AGRICULTURE.
XX
XX Theologis A, Sato T;
XX
XX WPI; 1991-368895/50.
XX
XX DNA encoding ACC synthase - used for control of plant development
XX and for prodn. of ACC synthase, ethylene and ethanol
XX
XX Disclosure: Fig 14; 73pp; English.
XX
XX Genomic clone LE-ACC1A was one of four different clones identified
XX in a library prepared from total DNA extracted from etiolated Rutgers
XX seedlings. Screening was by the cDNA clone pLACC1 (see AAQ15134) . The
XX amino acid sequence was deduced from the cDNA and shows
XX considerable homology with ACC sequences deduced from the three other
XX tomato and the zucchini clones.
XX See AAQ15131-Q15140.
XX
XX Sequence 469 AA:

Query Match          79.5%; Score 31; DB 12; Length 469;
Best Local Similarity 62.5%; Pred. No. 1-2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
Db 108 sfdsmnv 115

```

```

RESULT 13
AAW32619
ID AAW32619 standard; Protein: 964 AA.
XX
XX AAW32619;
AC
XX
DT 04-FEB-1998 (first entry)
XX
DE Cyclic-Isomaltoligosaccharide synthase.
XX
XX Cyclic-Isomaltoligosaccharide synthase; cycloisomaltotriose;
KM anti decay activity; Escherichia coli; E.coli; microorganism;
KW recombinant DNA.
XX
XX Bacillus sp.
OS
XX JF09234073-A.
PN
XX 09-SEP-1997.
PD
XX
XX 01-MAR-1996; 96JP-0045103.
PE
XX
XX 01-MAR-1996; 96JP-0045103.
PR
XX
XX (NORO ) NORINSUISANSO SHOKUHIN SOGO.
PA (NODA ) ZH NODA SANGYO KAGAKU KENKYUSHO.
XX
XX WPI; 1997-497314/46.
DR N-PSDB; AAT92376.
XX
XX Cyclic-iso:malt:oligosaccharide synthase - used in the production of
PT cyclo:iso:malt:heptose with strong anti-decay activity
XX
XX Claim 3; Page 8-10; 12pp; Japanese.
PS
XX
XX The present sequence represents a novel cyclic-Isomaltoligosaccharide
CC synthase isolated from a Bacillus sp. (possibly Bacillus circulans
CC U-155). The cyclic-Isomaltoligosaccharide synthase has the following
CC physicochemical properties: (1) action: acts on glucose polymers (e.g.
CC dextran) consisting of 1,6-alpha-linkages to form cyclic-
CC Isomaltoligosaccharides based on cycloisomaltotriose; (2) substrate
CC specificity: acts on dextran having 1,6-alpha-linkages as main chains,
CC but not on amylopectin, pullulan; and (3) optimum pH and stable pH
CC range: active around pH 6.0 and stable in the range of pH 4.5 to 9.0.
CC The DNA sequence may be used to produce cyclic-Isomaltoligosaccharide
CC synthase in microorganisms of the genus Escherichia, particularly
CC E.coli. The resultant cyclic-Isomaltoligosaccharide synthase can
CC efficiently produce cycloisomaltotriose which has very strong
CC anti-decay activity.
XX
XX Sequence 964 AA;
SQ
Query Match 79.5%; Score 31; DB 18; Length 964;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDSSN 6
DB 521 sydsn 526

```

```

KM Cell adhesion: intercellular adhesion molecule; endothelium;
KW Inflammation; leukocyte chemotaxis; rhinovirus; common cold.
XX
XX W09203473-A.
FN
XX
XX 05-MAR-1992.
PD
XX
XX 23-AUG-1991; 91WO-U006063.
PE
XX
XX 27-AUG-1990; 90US-0573624.
PR
XX
XX (CERTU ) CERTUS CORP.
PA
XX
XX Liu DY, Kaymakalan Z, Mundy K;
PL
XX
XX WPI; 1992-096833/12.
DR
XX
XX Peptide(s) derived from beta sub-unit CD18 of leukocyte integrins
PT - prevent leukocyte binding to ICAM and leukocyte chemotaxis, for
PT treating inflammatory diseases and rhinoviral infection
XX
XX Claim 2; Page 23; 31pp; English.
PS
XX
XX This peptide (derived from CD18 - see AAR24256) was synthesised and
CC tested for capacity to interfere with or block adhesion of
CC polymorphonuclear leukocytes to human endothelial cell monolayers.
CC The peptide shows inhibitory activity at 10(power -4)M. The peptide
CC can be used to treat/prevent inflammation and can inhibit binding
CC of virions to endothelium. See AAO22780 and AAR22104-R22112.
XX
XX Sequence 29 AA;
SQ
Query Match 76.9%; Score 30; DB 13; Length 29;
Best Local Similarity 87.5%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SYDSSNV 8
DB 10 sedssnv 17

```

```

RESULT 14
AAR22106
ID AAR22106 standard; Protein: 29 AA.
XX
XX AAR22106;
AC
XX
DT 15-JUL-1992 (first entry)
XX
DE Peptide 3-29 derived from CD18 leukocyte integrin.
XX

```

```

RESULT 15
AAW64528
ID AAW64528 standard; Protein: 126 AA.
XX
XX AAW64528;
AC
XX
DT 18-NOV-1998 (first entry)
XX
XX Papillomavirus major capsid protein L1 from plasmid DL369.
DE
XX
XX Major capsid protein L1; epithelial neoplasm; diagnosis; infection;
KW therapy; vaccine.
XX
XX Human papillomavirus.
OS
XX
XX Key Location/Qualifiers
FH Protein 1..126
FT /note="partial sequence"
XX
XX DE19735118-C1.
PN
XX
XX 13-AUG-1998.
PD
XX
XX 13-AUG-1997; 97DE-1035118.
PE
XX
XX 13-AUG-1997; 97DE-1035118.
PR
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
XX
XX De Villiers zur Hausen E, zur Hausen H;
PI
XX

```

DR WPI: 1998-415283/36.
 DR N-PSDB: AAY49545.
 XX
 XX
 PT Papilloma virus DNA - coding for major capsid protein
 XX
 PS Claim 1; Fig 3; 12p; German.
 XX
 CC This sequence represents a human papillomavirus major capsid protein L1
 CC which has been isolated from plasmid DL369 using total DNA from an
 CC epithelial neoplasm biopsy. Transformants can be cultured to produce
 CC the protein. The DNA can be used as a reagent for diagnosis of papilloma
 CC virus infections. The protein can be used for diagnosis or therapy of
 CC papilloma virus infections or for vaccination against such infections.
 XX
 S0 Sequence 126 AA;

Query Match 76.9%; Score 30; DB 19; Length 126;
 Best Local Similarity 83.3%; Pred. No. 44;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YDSSNV 7
 |||||:
 Db 30 ydsn1 35

Search completed: November 19, 2001, 08:12:37
 Job time: 90 sec


```
RESULT 2
US-09-610-118-66
; Sequence 66, Application US/09610118
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villaveal, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; APPLICANT: Gill, D.
; APPLICANT: Qian, M.
; APPLICANT: Kingsbury, G.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-211
; CURRENT APPLICATION NUMBER: US/09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2/14/00
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 12/6/99
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 6/30/99
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-610-118-66
```

```
Query Match
Best Local Similarity 100.0%; Score 39; DB 20; Length 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SYDSSNV 8
DB 1 SYDSSNV 8
```

```
RESULT 3
US-09-832-312-66
; Sequence 66, Application US/09832312
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-312-66
```

```
Query Match
Best Local Similarity 100.0%; Score 39; DB 22; Length 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SYDSSNV 8
DB 1 SYDSSNV 8
```

```
RESULT 4
PCT-US00-26524B-6775
; Sequence 6775, Application PC/TUS0026524B
; GENERAL INFORMATION:
; APPLICANT: Birse et. al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA003PCT
; CURRENT APPLICATION NUMBER: PCT/US00/26524B
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6775
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-26524B-6775
```

```
Query Match
Best Local Similarity 100.0%; Score 39; DB 1; Length 161;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SYDSSNV 8
DB 114 SYDSSNV 121
```

```
RESULT 5
US-60-324-109-30645
; Sequence 30645, Application US/60324109
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovallig, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)B
; CURRENT APPLICATION NUMBER: US/60/324,109
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 30645
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
US-60-324-109-30645
```

```
Query Match
Best Local Similarity 84.6%; Score 33; DB 24; Length 375;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SYDSSNV 8
DB 354 SYDSSNV 361
```

```
RESULT 6
US-60-173-464-19058
; Sequence 19058, Application US/60173464
; GENERAL INFORMATION:
```

```

; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: C1000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 19058
; LENGTH: 1945
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-19058

```

```

Query Match      84.6%; Score 33; DB 24; Length 1945;
Best Local Similarity 62.5%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYDSNNV 8
||:||||:
Db 1266 SYESSNII 1273

```

```

RESULT 7
US-60-191-637-23222
; Sequence 23222, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 23222
; LENGTH: 1945
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-23222

```

```

Query Match      84.6%; Score 33; DB 24; Length 1945;
Best Local Similarity 62.5%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYDSNNV 8
||:||||:
Db 1266 SYESSNII 1273

```

```

RESULT 8
US-60-191-681-18308
; Sequence 18308, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; FILE OF INVENTION: USES THEREOF.
; FILE REFERENCE: C1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 18308
; LENGTH: 1945
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-18308

```

```

Query Match      84.6%; Score 33; DB 24; Length 1945;
Best Local Similarity 62.5%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYDSNNV 8
||:||||:
Db 1266 SYESSNII 1273

```

```

RESULT 9
US-09-270-767-32375
; Sequence 32375, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32375
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32375

```

```

Query Match      82.1%; Score 32; DB 16; Length 96;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SYDSNNV 8
||:||||:
Db 85 SYDTANIV 92

```

```

RESULT 10
US-09-107-532-5317
; Sequence 5317, Application US/09107532
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
; NUMBER OF SEQUENCES: 7308
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085598
; FILING DATE: May 14, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007

```

TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5317:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...515
US-09-107-532-5317

Query Match 82.1%; Score 32; DB 15; Length 515;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDSSNV 7
:|||||
DB 175 AYDSSNV 181

RESULT 11
US-09-107-532A-5317
; Sequence 5317, Application US/09107532A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107.532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Atinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5317:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...515
; SEQUENCE DESCRIPTION: SEQ ID NO: 5317:

US-09-107-532A-5317

Query Match 82.1%; Score 32; DB 15; Length 515;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDSSNV 7
:|||||
DB 175 AYDSSNV 181

RESULT 12
US-60-167-217-9945
; Sequence 9945, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1000152
; CURRENT APPLICATION NUMBER: US/60/167,217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9945
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Drosophila
US-60-167-217-9945

Query Match 82.1%; Score 32; DB 24; Length 521;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDSSNV 8
:||||:|
DB 510 SYDTANIV 517

RESULT 13
US-60-191-637-9950
; Sequence 9950, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE REFERENCE: C1000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9950
; LENGTH: 521
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-9950

Query Match 82.1%; Score 32; DB 24; Length 521;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDSSNV 8
:||||:|
DB 510 SYDTANIV 517

RESULT 14
US-60-173-464-1878
; Sequence 1878, Application US/60173464

```
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ. ID NOS: 30269
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1878
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-1878
```

```
Query Match      82.1%; Score 32; DB 24; Length 542;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 SYDSSNV 7
        :|||||
Db       95 NYDSSNV 101
```

```
RESULT 15
US-60-191-637-2271
; Sequence 2271, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2271
; LENGTH: 342
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-2271
```

```
Query Match      82.1%; Score 32; DB 24; Length 542;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 SYDSSNV 7
        :|||||
Db       95 NYDSSNV 101
```

Search completed: November 19, 2001, 08:23:17
Job time: 730 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:23:32 ; Search time 11.44 Seconds
(without alignments)
3.725 Million cell updates/sec

Title: US-09-610-118-66

Perfect score: 39

Sequence: 1 SYDSSNV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 25229 seqs, 5326477 residues

Total number of hits satisfying chosen parameters: 25229

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCIT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEM_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US08_NEM_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEM_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEM_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEM_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	79.5	10	US-09-972-656-23	Sequence 23, Appl
2	31	79.5	103	US-09-972-656-130	Sequence 130, Appl
3	31	79.5	217	US-09-972-656-88	Sequence 88, Appl
4	29	74.4	804	US-09-815-242-10393	Sequence 10393, A
5	29	74.4	804	US-09-815-242-103920	Sequence 103920, A
6	28	71.8	11	US-09-972-656-31	Sequence 31, Appl
7	28	71.8	218	US-09-972-656-102	Sequence 102, Appl
8	28	71.8	315	US-09-886-055-445	Sequence 445, Appl
9	28	71.8	316	US-09-886-055-133	Sequence 133, Appl
10	28	71.8	350	US-09-815-242-14039	Sequence 14039, A
11	28	71.8	638	US-09-815-242-5222	Sequence 5222, A
12	28	71.8	642	US-09-815-242-12143	Sequence 12143, A
13	28	71.8	798	US-09-978-249-18	Sequence 48, Appl
14	27	69.2	17	US-09-972-656-48	Sequence 102, Appl
15	27	69.2	225	US-09-453-234-102	Sequence 108, Appl
16	27	69.2	225	US-09-453-234-108	Sequence 110, Appl
17	27	69.2	225	US-09-453-234-110	Sequence 110, Appl
18	27	69.2	227	US-09-972-656-76	Sequence 16, Appl
19	27	69.2	372	US-60-325-537-15	Sequence 15, Appl
20	27	69.2	592	US-09-850-716A-169	Sequence 169, Appl
21	27	69.2	791	US-09-850-716A-170	Sequence 170, Appl
22	27	69.2	881	US-09-850-716A-430	Sequence 430, Appl
23	27	69.2	920	US-09-850-716A-357	Sequence 357, Appl
24	27	69.2	942	US-09-974-298-71	Sequence 71, Appl
25	27	69.2	943	US-09-850-716A-161	Sequence 161, Appl
26	27	69.2	1930	US-60-325-427-2	Sequence 2, Appl
27	26	66.7	312	US-09-779-679-35	Sequence 35, Appl

28	26	66.7	396	5	US-09-815-242-11791	Sequence 11791, A
29	26	66.7	352	5	US-09-815-242-14061	Sequence 14061, A
30	26	66.7	804	5	US-09-897-516-5911	Sequence 5911, A
31	26	66.7	806	5	US-09-815-242-11774	Sequence 11774, A
32	26	66.7	855	5	US-09-644-600-2	Sequence 2, Appl
33	26	66.7	863	5	US-09-976-594-10	Sequence 10, Appl
34	26	66.7	1178	5	US-09-851-194-2	Sequence 2, Appl
35	26	66.7	1599	5	US-09-897-516-4687	Sequence 4687, A
36	25	64.1	79	5	US-09-815-242-5861	Sequence 5861, A
37	25	64.1	79	5	US-09-815-242-12945	Sequence 12945, A
38	25	64.1	79	5	US-09-815-242-13095	Sequence 13095, A
39	25	64.1	136	5	US-09-897-516-6168	Sequence 6168, A
40	25	64.1	186	5	US-09-897-516-4802	Sequence 4802, A
41	25	64.1	236	5	US-09-576-424-10	Sequence 10, Appl
42	25	64.1	243	5	US-09-966-521-70	Sequence 70, Appl
43	25	64.1	243	5	US-09-966-521-2	Sequence 2, Appl
44	25	64.1	337	5	US-09-615-846A-19	Sequence 19, Appl
45	25	64.1	354	5	US-09-615-846A-21	Sequence 21, Appl

ALIGNMENTS

```
RESULT 1
US-09-972-656-23
; Sequence 23, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972.656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-656-23

Query Match      79.5%; Score 31; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYDSSN 6
DB      2 SYDSSN 7

RESULT 2
US-09-972-656-130
; Sequence 130, Application US/09972656
; GENERAL INFORMATION:
; APPLICANT: Deshpande, Rajendra
; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
; FILE REFERENCE: A-799
; CURRENT APPLICATION NUMBER: US/09/972.656
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC.
; LOCATION: (7)..(7)
; OTHER INFORMATION: Unidentified
```

```
NAME/KEY: Misc.
LOCATION: (33)..(33)
OTHER INFORMATION: Unidentified
NAME/KEY: Misc.
LOCATION: (34)..(34)
OTHER INFORMATION: Unidentified
NAME/KEY: Misc.
LOCATION: (35)..(35)
OTHER INFORMATION: Unidentified
NAME/KEY: Misc.
LOCATION: (36)..(36)
OTHER INFORMATION: Unidentified
US-09-972-656-130
```

```
Query Match
Best Local Similarity 79.5%; Score 31; DB 5; Length 103;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SYDSSN 6
DB 98 SYDSSN 103
```

```
RESULT 3
US-09-972-656-88
Sequence 88, Application US/09972656
GENERAL INFORMATION:
APPLICANT: Deshpande, Rajendra
TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
FILE REFERENCE: A-799
CURRENT APPLICATION NUMBER: US/09/972,656
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn version 3.0
SEQ ID NO 88
LENGTH: 217
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-656-88
```

```
Query Match
Best Local Similarity 79.5%; Score 31; DB 5; Length 217;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SYDSSN 6
DB 93 SYDSSN 98
```

```
RESULT 4
US-09-815-242-10393
Sequence 10393, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
```

```
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10393
LENGTH: 804
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10393
```

```
Query Match
Best Local Similarity 74.4%; Score 29; DB 5; Length 804;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SYDSSNV 7
DB 4 SYDSSSI 10
```

```
RESULT 5
US-09-815-242-13920
Sequence 13920, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13920
LENGTH: 804
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-13920
```

```
Query Match
Best Local Similarity 74.4%; Score 29; DB 5; Length 804;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 SYDSSNV 7
 11111:
 DB 4 SYDSSSI 10

RESULT 6

US-09-972-656-31
 ; Sequence 31, Application US/09972656
 ; GENERAL INFORMATION:
 ; APPLICANT: Deshpande, Rajendra
 ; APPLICANT: Tsai, Mei-Mei
 ; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
 ; FILE REFERENCE: A-799
 ; CURRENT APPLICATION NUMBER: US/09/972,656
 ; CURRENT FILING DATE: 2001-10-05
 ; NUMBER OF SEQ ID NOS: 135
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 31
 ; LENGTH: 11
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-972-656-31

Query Match 71.8%; Score 28; DB 5; Length 11;
 Best Local Similarity 83.3%; Pred. No. 0.64;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 6
 11111:
 DB 2 SYDSSNV 7

RESULT 7

US-09-972-656-102
 ; Sequence 102, Application US/09972656
 ; GENERAL INFORMATION:
 ; APPLICANT: Deshpande, Rajendra
 ; APPLICANT: Tsai, Mei-Mei
 ; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
 ; FILE REFERENCE: A-799
 ; CURRENT APPLICATION NUMBER: US/09/972,656
 ; CURRENT FILING DATE: 2001-10-05
 ; NUMBER OF SEQ ID NOS: 135
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 102
 ; LENGTH: 218
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-972-656-102

Query Match 71.8%; Score 28; DB 5; Length 218;
 Best Local Similarity 83.3%; Pred. No. 16;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 6
 11111:
 DB 93 SYDSSNV 98

RESULT 8

US-09-886-055-445
 ; Sequence 445, Application US/09886055
 ; GENERAL INFORMATION:
 ; APPLICANT: STRAYER, LUBERT
 ; APPLICANT: ZOZULYA, SERGEY
 ; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 ; FILE REFERENCE: 078003-0277150
 ; CURRENT APPLICATION NUMBER: US/09/886,055

; CURRENT FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 60/213,812
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 522
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 445
 ; LENGTH: 315
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-886-055-445

Query Match 71.8%; Score 28; DB 5; Length 315;
 Best Local Similarity 62.5%; Pred. No. 24;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSSNV 8
 11111:
 DB 167 SYDSSNV 174

RESULT 9

US-09-886-055-133
 ; Sequence 133, Application US/09886055
 ; GENERAL INFORMATION:
 ; APPLICANT: STRAYER, LUBERT
 ; APPLICANT: ZOZULYA, SERGEY
 ; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 ; FILE REFERENCE: 078003-0277150
 ; CURRENT APPLICATION NUMBER: US/09/886,055
 ; CURRENT FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 60/213,812
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 522
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 133
 ; LENGTH: 316
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-886-055-133

Query Match 71.8%; Score 28; DB 5; Length 316;
 Best Local Similarity 62.5%; Pred. No. 24;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSSNV 8
 11111:
 DB 167 SYDSSNV 174

RESULT 10

US-09-815-242-14039
 ; Sequence 14039, Application US/09815242
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23

```

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14039
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14039
```

```
Query Match          71.8%; Score 28; DB 5; Length 350;
Best Local Similarity 71.4%; Pred. No. 27;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SYDSNV 7
        1:1:111
Db       126 SFDASNV 132
```

```

RESULT 11
; Sequence 5222, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 60/03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5222
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5222
```

```
Query Match          71.8%; Score 28; DB 5; Length 638;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 SYDSNV 7
```

```
Db       54 NYDSGNV 60
```

```

RESULT 12
; Sequence 12143, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 60/03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12143
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12143
```

```
Query Match          71.8%; Score 28; DB 5; Length 642;
Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 SYDSNV 7
        1:11 11
Db       54 NYDSGNV 60
```

```

RESULT 13
; Sequence 8, Application US/09978249
; GENERAL INFORMATION:
; APPLICANT: Fiscella, et al.
; TITLE OF INVENTION: Extracellular Matrix Polynucleotides, Polypeptides, and Antibio
; FILE REFERENCE: PT054PI
; CURRENT APPLICATION NUMBER: US/09/978,249
; PRIOR APPLICATION NUMBER: 2001-10-17
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: PCT/US01/11643
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-249-8
```

Query Match 71.8%; Score 28; DB 5; Length 798;
 Best Local Similarity 75.0%; Pred. No. 65;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Search completed: November 19, 2001, 08:23:32
 Job time: 744 sec

OY 1 SYDSSNV 8
 DB 360 SMDSSNVL 367

RESULT 14
 US-09-972-656-48
 ; Sequence 48, Application US/09972656
 ; GENERAL INFORMATION:
 ; APPLICANT: Deshpande, Rajendra
 ; APPLICANT: Tsai, Mei-Mei
 ; TITLE OF INVENTION: Fully Human Antibody Fab Fragments with Human Interferon-Gamma
 ; FILE REFERENCE: A-799
 ; CURRENT APPLICATION NUMBER: US/09/972,656
 ; CURRENT FILING DATE: 2001-10-05
 ; NUMBER OF SEQ ID NOS: 135
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 48
 ; LENGTH: 17
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-972-656-48

Query Match 69.2%; Score 27; DB 5; Length 17;
 Best Local Similarity 83.3%; Pred. No. 1.7;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSSN 6
 DB 3 SYDGSN 8

RESULT 15
 US-09-453-234-102
 ; Sequence 102, Application US/09453234
 ; GENERAL INFORMATION:
 ; APPLICANT: Buechler, Joe
 ; APPLICANT: Walkers, Gunars
 ; APPLICANT: Gray, Jeff
 ; APPLICANT: Lomberg, Nils
 ; APPLICANT: Biosite Diagnostics, Inc.
 ; TITLE OF INVENTION: Genpharm International
 ; FILE REFERENCE: 020015-000110US
 ; CURRENT APPLICATION NUMBER: US/09/453,234
 ; CURRENT FILING DATE: 1999-12-01
 ; PRIOR APPLICATION NUMBER: US 60/157,415
 ; PRIOR FILING DATE: 1999-10-02
 ; NUMBER OF SEQ ID NOS: 112
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 102
 ; LENGTH: 225
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; OTHER INFORMATION: M2-31H
 US-09-453-234-102

Query Match 69.2%; Score 27; DB 5; Length 225;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYDSSN 6
 DB 52 SYDGSN 57

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2001, 08:13:26 ; Search time 43.5 Seconds
(without alignments)
4.139 Million cell updates/sec

Title: US-09-610-118-66
39
Sequence: 1 SYDSSNVV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	31	79.5	222	1 US-07-869-933-12	Sequence 12, Appl
2	31	79.5	222	1 US-07-869-933-28	Sequence 28, Appl
3	31	79.5	222	4 US-09-103-663-12	Sequence 12, Appl
4	31	79.5	222	4 US-09-103-663-28	Sequence 28, Appl
5	30	76.9	222	5 PCT-US96-01314-58	Sequence 58, Appl
6	30	76.9	223	5 PCT-US96-01314-53	Sequence 53, Appl
7	30	76.9	618	4 US-08-961-083-72	Sequence 72, Appl
8	30	76.9	676	3 US-08-630-172-10	Sequence 10, Appl
9	30	76.9	676	4 US-09-375-419-10	Sequence 10, Appl
10	30	76.9	769	2 US-08-789-078-1	Sequence 1, Appl
11	30	76.9	769	2 US-08-752-633-1	Sequence 1, Appl
12	30	76.9	769	2 US-08-476-062A-45	Sequence 45, Appl
13	30	76.9	769	2 US-07-728-215-31	Sequence 31, Appl
14	30	76.9	769	5 PCT-US95-04886-1	Sequence 1, Appl
15	30	76.9	769	5 PCT-US96-01314-45	Sequence 45, Appl
16	29	74.4	16	2 US-08-932-876-1	Sequence 1, Appl
17	29	74.4	25	1 US-08-378-761A-30	Sequence 30, Appl
18	29	74.4	25	1 US-08-485-286-30	Sequence 30, Appl
19	29	74.4	25	6 5248606-16	Patent No. 5248606
20	29	74.4	40	2 US-08-679-405-6	Sequence 6, Appl
21	29	74.4	40	2 US-08-842-799-6	Sequence 6, Appl
22	29	74.4	40	5 PCT-US96-11458-6	Sequence 6, Appl
23	29	74.4	109	4 US-09-157-370-5	Sequence 5, Appl
24	29	74.4	459	4 US-08-378-313-33	Sequence 33, Appl
25	29	74.4	600	2 US-08-679-405-2	Sequence 2, Appl
26	29	74.4	600	2 US-08-842-799-2	Sequence 2, Appl
27	29	74.4	600	4 US-09-271-778-2	Sequence 2, Appl

28	29	74.4	600	5 PCT-US96-11458-2	Sequence 2, Appl
29	29	74.4	740	2 US-08-410-784A-4	Sequence 4, Appl
30	29	74.4	818	2 US-08-410-784A-2	Sequence 2, Appl
31	29	74.4	818	4 US-09-346-237-11	Sequence 11, Appl
32	28	71.8	18	1 US-07-800-364B-2	Sequence 2, Appl
33	28	71.8	18	5 PCT-US91-03388-2	Sequence 2, Appl
34	28	71.8	49	1 US-07-800-364B-8	Sequence 8, Appl
35	28	71.8	49	5 PCT-US91-03388-8	Sequence 8, Appl
36	28	71.8	102	3 US-08-478-097A-17	Sequence 17, Appl
37	28	71.8	102	4 US-08-931-858E-162	Sequence 162, App
38	28	71.8	102	4 US-08-981-739-162	Sequence 162, App
39	28	71.8	112	1 US-07-800-364B-15	Sequence 15, Appl
40	28	71.8	112	5 PCT-US91-03388-13	Sequence 13, Appl
41	28	71.8	119	4 US-08-624-635-15	Sequence 15, Appl
42	28	71.8	139	1 US-08-278-729A-7	Sequence 7, Appl
43	28	71.8	139	1 US-08-278-729A-8	Sequence 8, Appl
44	28	71.8	139	1 US-08-155-343A-7	Sequence 7, Appl
45	28	71.8	139	1 US-08-155-343A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-07-869-933-12
Sequence 12, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869, 933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Rat
STRAIN: PCRI alpha subunit
US-07-869-933-12

Query Match 79.5%; Score 31; DB 1; Length 222;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
||||:|:
Db 130 SYDSSNI 136

RESULT 2
US-07-869-933-28
; Sequence 28, Application US/07869933
; Patent No. 5770396

GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN E.
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22113-0299

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NHHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149

INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: alpha subunit
; US-07-869-933-28

Query Match 79.5%; Score 31; DB 1; Length 222;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
||||:|:
Db 130 SYDSSNI 136

RESULT 3
US-09-103-663-12
; Sequence 12, Application US/09103663D
; Patent No. 6171803

GENERAL INFORMATION:
; APPLICANT: Kinet et al.
; TITLE OF INVENTION: Isolation, characterization, and use of the human beta
; TITLE OF INVENTION: subunit of the high affinity receptor for
; FILE REFERENCE: 50490
; CURRENT APPLICATION NUMBER: US/09/103,663D
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 07/869,933
; EARLIER FILING DATE: 1992-04-16

; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-103-663-12

Query Match 79.5%; Score 31; DB 4; Length 222;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
||||:|:
Db 130 SYDSSNI 136

RESULT 4
US-09-103-663-28
; Sequence 28, Application US/09103663D
; Patent No. 6171803

GENERAL INFORMATION:
; APPLICANT: Kinet et al.
; TITLE OF INVENTION: Isolation, characterization, and use of the human beta
; TITLE OF INVENTION: subunit of the high affinity receptor for
; FILE REFERENCE: 50490
; CURRENT APPLICATION NUMBER: US/09/103,663D
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 07/869,933
; EARLIER FILING DATE: 1992-04-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-103-663-28

Query Match 79.5%; Score 31; DB 4; Length 222;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDSSNV 7
||||:|:
Db 130 SYDSSNI 136

RESULT 5
PCT-US96-01314-58
; Sequence 58, Application PC/TUS9601314
; GENERAL INFORMATION:

APPLICANT: M. Amin Arnaut
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-58

Query Match 76.9%; Score 30; DB 5; Length 252;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYDSSNV 8
11111111
Db 237 SEDSSNV 244

RESULT 6
PCT-US96-01314-53
Sequence 53, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-53

Query Match 76.9%; Score 30; DB 5; Length 253;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYDSSNV 8
11111111
Db 238 SEDSSNV 245

RESULT 7
US-08-961-083-72
Sequence 72, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS Version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 618 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-72

Query Match 76.9%; Score 30; DB 4; Length 618;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYDSSNV 8
11111111
Db 363 SYDSSNV 370

RESULT 8
US-08-630-172-10
Sequence 10, Application US/08630172
Patent No. 6060054
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,172
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 676 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-172-10

Query Match 76.9% Score 30; DB 3; Length 676;
Best Local Similarity 87.5% Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STDSSNV 8
I I I I I I I I
DB 324 SEDSSNV 331

RESULT 9
US-09-375-419-10
Sequence 10, Application US/09375419
Patent No. 6264950
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/375,419
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,172
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 676 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-375-419-10

Query Match 76.9% Score 30; DB 4; Length 676;
Best Local Similarity 87.5% Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STDSSNV 8
I I I I I I I I
DB 324 SEDSSNV 331

RESULT 10
US-08-789-078-1
Sequence 1, Application US/08789078
Patent No. 5843885
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Slabean, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbels, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789,078
FILING DATE: 03-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,513
FILING DATE: 19-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 22833
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: (816)474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label= signal

```

: OTHER INFORMATION: /note= "signal sequence"
: FEATURE:
: NAME/KEY: Region
: LOCATION: 449..496
: OTHER INFORMATION: /label= repeat
: OTHER INFORMATION: /note= "cysteine rich repeat"
: FEATURE:
: NAME/KEY: Region
: LOCATION: 497..540
: OTHER INFORMATION: /label= repeat
: OTHER INFORMATION: /note= "cysteine rich repeat"
: FEATURE:
: NAME/KEY: Region
: LOCATION: 541..581
: OTHER INFORMATION: /label= repeat
: OTHER INFORMATION: /note= "cysteine rich repeat"
: FEATURE:
: NAME/KEY: Region
: LOCATION: 582..617
: OTHER INFORMATION: /label= repeat
: OTHER INFORMATION: /note= "cysteine rich repeat"
: FEATURE:
: NAME/KEY: Domain
: LOCATION: 701..723
: OTHER INFORMATION: /label= trans
: OTHER INFORMATION: /note= "transmembrane domain"
: FEATURE:
: NAME/KEY: Domain
: LOCATION: 724..769
: OTHER INFORMATION: /label= cyto
: OTHER INFORMATION: /note= "cytoplasmic domain"
: PUBLICATION INFORMATION:
: AUTHORS: Pigott,
: TITLE: LFA-1 Amino acid sequence (B2) (from human
: TITLE: tonsil)
: JOURNAL: The Adhesion Molecule Facts Book
: PAGES: 96-96
: DATE: 1993
: RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
: US-08-789-078-1

Query Match      76.9%; Score 30; DB 2; Length 769;
Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SYDSSNV 8
Db      346 SEDSSNV 353

RESULT 11
US-08-752-633-1
: Sequence 1, Application US/08752633
: Patent No. 5863889
: GENERAL INFORMATION:
: APPLICANT: Benedict, Stephen H.
: APPLICANT: Sahaan, Teruna
: APPLICANT: Chan, Marcia
: APPLICANT: Tibbets, Scott
: TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
: TITLE OF INVENTION: IMMUNE TOLERANCE
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
: STREET: 1101 Walnut St.
: CITY: Kansas City
: STATE: MO
: COUNTRY: USA
: ZIP: 64106
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/752.633
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Collins, John M.
: REGISTRATION NUMBER: 26262
: REFERENCE/DOCKET NUMBER: 22833
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (816)474-9050
: TELEFAX: 434-363
: TELETYPE: 434-363
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 769 amino acids
: TYPE: amino acid
: TOPOLOGY: both
: MOLECULE TYPE: Protein
: HYPOTHEICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: TISSUE TYPE: Tonsil
: FEATURE:
: NAME/KEY: Region
: LOCATION: 1..22
: OTHER INFORMATION: /label= signal
: OTHER INFORMATION: /note= "signal sequence"
: FEATURE:
: NAME/KEY: Region
: LOCATION: 449..496
: OTHER INFORMATION: /label= repeat
: OTHER INFORMATION: /note= "cysteine rich repeat"
: FEATURE:
: NAME/KEY: Region
: LOCATION: 497..540
: OTHER INFORMATION: /label= repeat
: OTHER INFORMATION: /note= "cysteine rich repeat"
: FEATURE:
: NAME/KEY: Region
: LOCATION: 541..581
: OTHER INFORMATION: /label= repeat
: OTHER INFORMATION: /note= "cysteine rich repeat"
: FEATURE:
: NAME/KEY: Domain
: LOCATION: 701..723
: OTHER INFORMATION: /label= trans
: OTHER INFORMATION: /note= "transmembrane domain"
: FEATURE:
: NAME/KEY: Domain
: LOCATION: 724..769
: OTHER INFORMATION: /label= cyto
: OTHER INFORMATION: /note= "cytoplasmic domain"
: PUBLICATION INFORMATION:
: AUTHORS: Pigott,
: TITLE: LFA-1 Amino acid sequence (B2) (from human
: TITLE: tonsil)
: JOURNAL: The Adhesion Molecule Facts Book
: PAGES: 96-96
: DATE: 1993
: RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
: US-08-752-633-1

Query Match      76.9%; Score 30; DB 2; Length 769;
```

Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSNVV 8
1 | 111111
DB 346 SEDSNVV 353

RESULT 12
US-08-476-062A-45
; Sequence 45, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arrabout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-476-062A-45

Query Match 76.9%; Score 30; DB 2; Length 769;
Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSNVV 8
1 | 111111
DB 346 SEDSNVV 353

RESULT 13
US-07-728-215-31
; Sequence 31, Application US/07728215
; Patent No. 5962643
; GENERAL INFORMATION:

APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pyzela, Robert
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-728-215-31

Query Match 76.9%; Score 30; DB 2; Length 769;
Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSNVV 8
1 | 111111
DB 346 SEDSNVV 353

RESULT 14
PCT-US95-04886-1
; Sequence 1, Application PC/TUS9504886
; GENERAL INFORMATION:
; APPLICANT: Benedict, Stephen H.
; APPLICANT: Slahaan, Teruna
; APPLICANT: Chan, Marcia
; APPLICANT: Tibbels, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
TITLE OF INVENTION: IMMUNE TOLERANCE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04886
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 22833
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: 816)474-9057
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label= signal
FEATURE:
NAME/KEY: Region
LOCATION: 449..496
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 582..617
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label= trans
OTHER INFORMATION: /note= "transmembrane domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label= cyto
OTHER INFORMATION: /note= "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
TITLE: LFA-1 Amino acid sequence (B2) (from human
TITLE: tonsil)
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
PCT-US95-04886-1
Query Match 76.9%; Score 30; DB 5; Length 769;
Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYDSSNV 8
1 |||||
DB 346 SEDSSNV 353

RESULT 15
PCT-US96-01314-45
Sequence 45, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaut
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 769
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-45

Query Match 76.9%; Score 30; DB 5; Length 769;
Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SYDSSNV 8
1 |||||
DB 346 SEDSSNV 353

Search completed: November 19, 2001, 08:13:27
Job time: 140 sec

